

ARI Working Papers

Fort Bliss Field Unit

1986-1988

**Reproduced From
Best Available Copy**

These working papers are published in order to archive material that was not included in other ARI publications. The material contained herein may not meet ARI's usual scientific or professional standards for publication.

July 2001

United States Army Research Institute for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited.

20011102 061

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) July 2001		2. REPORT TYPE Final		3. DATES COVERED (from. . . to) 1986-1988	
4. TITLE AND SUBTITLE ARI Working Papers: Fort Bliss Field Unit, 1986-1988				5a. CONTRACT OR GRANT NUMBER	
				5b. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Allender, L.E., Barber, A.V., Drewfs, P.R., Lockhart, J.M., Bittner, A.C., Zaklad, A., Dick, A.O., Herman, E.D., dePontbriand, R.J., Frederickson, E.W., Glenn, F.A. III, Wierwille, W.W., Iavecchia, H.P., Wherry, R.J. Jr., Linton, P.M., Hill, S.G., Lysaght, R.J., Bulger				5c. PROJECT NUMBER	
				5d. TASK NUMBER	
				5e. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: TAPC-ARI-PO 5001 Eisenhower Avenue Alexandria, VA 22333-5600				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue Alexandria, VA 22333-5600				10. MONITOR ACRONYM ARI	
				11. MONITOR REPORT NUMBER WP Bliss	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES ARI working papers were originally unofficial documents intended for limited distribution to obtain comments. These working papers are being archived in order to preserve material that was not included in other ARI publications. The material contained herein may not meet ARI's usual scientific or professional standards for publication.					
14. ABSTRACT (<i>Maximum 200 words</i>): Eight working papers dealing with the Patriot system, the Stinger, RADES, and operator workload .					
15. SUBJECT TERMS Patriot system, Stinger, RADES, operator workload, OWL, MANPRINT					
SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT Unlimited	20. NUMBER OF PAGES 563	21. RESPONSIBLE PERSON (Name and Telephone Number) David W. Witter (703) 617-0324
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified			

Fort Bliss Field Unit Working Papers

Allender, L.E. (1986). Preliminary recommendations on the use of color in the Patriot display, WP BLISS 86-01.

Allender, L.E. (1987). An evaluation of the usefulness of color in the Patriot display. WP BLISS 87-01.

Barber, A.V., Drewfs, P.R., & Johnson, D.M. (1987). Performance of Stinger teams using the RADES multiple weapon configuration. WP BLISS 87-09.

Barber, A.V., Drewfs, P.R., & Lockhart, J.M. (1987). Effective Stinger training in RADES. WP BLISS 87-02.

Bittner, A.C., Zaklad, A., Dick, A.O., & Wherry, R.J. (1987). Operator Workload (OWL) assessment program for the Army: Validation and analysis plans for three systems. WP BLISS 87-08.

dePontbriand, R.J., & Frederickson, E.W. (1988). Pedestal mounted Stinger candidate evaluation: Overview of MANPRINT data. WP BLISS 88-01.

Glenn, F.A. III, Bittner, A.C. Jr., Wierwille, W.W., Iavecchia, H.P., Wherry, R.J. Jr., & Linton, P.M. (1987). Foundations for an Operator Workload (OWL) assessment program for the Army. WP BLISS 87-05.

Hill, S.G., Lysaght, R.J., Bittner, A.C., & Bulger, J. (1987). Operator Workload (OWL) assessment program for the Army: Results from requirements document review and user interview analysis. WP BLISS 87-06.

[Handwritten scribble]

Working Paper

FB 86-01

PRELIMINARY RECOMMENDATIONS ON THE USE OF
COLOR IN THE PATRIOT DISPLAY

LAUREL E. ALLENDER



**U.S. Army Research Institute
for the Behavioral and Social Sciences**
5001 Eisenhower Avenue, Alexandria VA 22333

This working paper is an unofficial document intended for limited distribution to obtain comments. The views, opinions, and/or findings contained in this document are those of the author(s) and should not be construed as the official position of ARI or as an official Department of the Army position, policy, or decision, unless so designated by other official documentation.

**Reproduced From
Best Available Copy**

FACT SHEET

PERI-SB

15 May 86

SUBJECT: Preliminary Recommendations on the Use of Color in the Patriot Display

FACTS:

1. The ARI Fort Bliss Field Unit and DCD, USAADASCH are evaluating how well the use of color on the Patriot display contributes to overall system performance by improving operator performance. Two developments prompted this effort: (1) Color display technology has advanced to the point where literally millions of colors can be generated and displayed in high resolution, real time graphics. (2) Patriot operators are being presented with rapidly changing, high density information on a complex display that pushes and sometimes exceeds their information processing limits.

2. Color Considerations. A survey of background research on color considerations was conducted. Key considerations for display design and operator performance are as follows: No more than 8 to 15 different colors should be used in a single display and those colors must be chosen carefully to ensure sufficient color contrast. Existing military color-meaning associations should be used; however, new color-meaning pairings can be learned readily. Finally, although color has been proven to be useful for certain specific tasks, this usefulness does not always translate into an improvement in overall performance of a complex task.

3. Preliminary Recommendations on Color Use. Based on the color considerations, an experimental display was developed and the specifics are presented as preliminary recommendations. The equipment is a Raster Technologies Model One/80 color display which was integrated with the Patriot Tactical Operations Simulator at DCD. The overall color scheme for the situation display area of the Patriot display is designed to reduce clutter and promote the visual grouping of similar objects. Volumes are drawn as filled polygons. Military color-meaning associations are the basis for the color assignments and objects that provide similar information are assigned the same color. To ensure sufficient contrast, background objects are of low brightness and saturation whereas foreground objects are of high brightness and saturation. The alert message line and tabular display area both have light gray backgrounds and black text. Color coding is used on the Engaged Data tabular display to alert the operator when tracks have met certain critical criteria.

4. Color Display Study. A study to compare the performance of trained Patriot operators using the experimental display with their performance using the current monochrome display is underway. Testing of 24 operators from 1ST BN, 43D ADA and seven operators from the IFFN/JTF began 3 March 1986 and is scheduled for completion 21 May 1986. Data analysis is underway and a final report of the results is due 15 September 1986.

Strub
Chief
8-4491

PRELIMINARY RECOMMENDATIONS ON THE USE OF COLOR IN THE PATRIOT DISPLAY

- WORKING PAPER -

Ms. Laurel Allender, Research Psychologist

US Army Research Institute for the Behavioral and Social Sciences

Fort Bliss Field Unit

15 May 1986

TABLE OF CONTENTS

	Page
Introduction	1
Color Considerations	1
Color Display System	2
Patriot Display	2
Recommendations	3
Summary Comments	11
References	12
Appendix A Color Display Study	13
Appendix B Design Considerations	16

LIST OF TABLES AND FIGURES

	Page
Figure 1 Patriot Display	4
Table 1 Situation Display Objects	5-6
Figure 2 Alert Message Line	7
Figure 3 Tabular Display Area	7

The recommendations on the use of color in the Patriot display presented here are preliminary recommendations based on a survey of color research and on the experience of developing an experimental display. The color research survey and the color display development are part of a joint effort between the Army Research Institute (ARI) Fort Bliss Field Unit and the Directorate of Combat Developments (DCD), US Army Air Defense Artillery School (USAADASCH) to investigate whether using color contributes to overall system performance by improving the operator's ability to understand and interact with the system. A third component of the effort is a study to compare the performance of trained Patriot operators using the experimental color display with their performance using existing monochrome display. Appendix A includes an outline of the color display study currently being conducted by ARI.

Two developments have prompted this investigation of the potential uses of color on the Patriot display. One development is that color display technology has advanced to the point where literally millions of colors can be generated and displayed in high resolution, real time graphics. Formerly, technology limited the use of color to outlined symbols or text in only three or four different colors. The second development is a direct result of advances in computer technology generally: Since computers are now faster and capable of displaying more information than ever before, the information processing limits of the human are being pushed and sometimes exceeded. This is true with respect to the Patriot where the operators are presented with rapidly changing, high density information on a complex display. It is possible for the human to adapt to this increased load by developing alternative information processing and decision strategies. Unfortunately, such strategies are all too likely to be shortcuts and therefore unsatisfactory. More satisfactorily, the computer can be adapted by further improving the display of information.

Color Considerations

Many of the techniques shown by psychological research to improve the display of information are already employed in the design of the Patriot display -- symbol, shape and line coding, brightness coding, blink-rate coding, and formatting of tables. Another known technique is the use of color for coding, highlighting, or background shading. In the application of color, more so than in the application of other display enhancing techniques, there are many factors to be considered--perceptual, psychological and practical factors. A complete discussion of these factors is presented in Appendix B, but a few will be discussed briefly here.

To begin with, a basic perceptual factor is that color is actually multidimensional, varying along the dimensions of hue, brightness and saturation (i.e., degree of whiteness) although people perceive it unidimensionally. People can distinguish up to 150 colors in side by side comparisons, but can consistently identify or name only about 8 to 15. Further, colors are not all brought into focus equally well nor are they perceived as being of equal brightness. A last perceptual factor, but one with serious implications for computer display design, is that the perception

of the color of an object changes depending on the color of the surrounding area or background, particularly if the object is small.

Of the psychological factors, the most obvious, but perhaps the most misunderstood is that of color-meaning associations. There are some widely shared associations, green with go, yellow with caution, and red with stop or danger. However, most color-meaning associations are fairly context-specific and can be readily developed through practice. In the military, green or blue have become associated with the concept friendly, yellow or white with unknown, and red with hostile. Another important psychological factor is that color increases the amount of information that can be seen with a single glance because it can be perceived before conscious attention is paid to it. Thus, color is useful for warning signals and when searching for color coded information. It also enhances the grouping of similar objects and at the same time reduces the clutter on high density displays.

Finally, the practical application of color to other displays has demonstrated that there are tradeoffs and limitations in using color: It cannot code background and foreground information or different sections of a display with equal priority. It cannot be used effectively to code categories of information on more than one level. For example, it cannot be used effectively to code both the identity and the altitude of an aircraft. Use of color has been shown to enhance performance of some tasks such as obtaining information from a display; however, it does not necessarily enhance performance of more complex tasks such as landing an aircraft. A final practical factor is that there is some evidence that color may reduce subjective workload, that is, how hard a task seems, so that there may be an indirect enhancement of performance even if there is not a direct one.

Color Display System

The recommendations on the use of color are made with respect to the specific color display system obtained for this project, a Raster Technologies Model One/80. The key capabilities of the Model One/80 are real time interactive graphics, 1280 x 1024 pixel (picture element) display resolution, a palette of 16.7 million colors, multiple display windows, and polygon fill. This latter capability means that hostile aircraft tracks can still be outlined diamonds as on the current display, but also that areas and volumes can be solidly filled polygons rather than dashed or outlined. Although 16.7 million different colors can be generated, only a subset of that number can be displayed simultaneously given the other display requirements. So instead of a simultaneous display capability, the 16.7 million colors provide needed flexibility in selecting precise shades. The experimental display configuration displays a subset of 27 different colors.

Patriot Display

It will be useful before discussing the recommendations to describe the Patriot display. The description is of the Engagement Control Station (ECS) at the fire platoon. Much of the description will also apply to the Information Coordination Central (ICC) at the battalion; however, some aspects of the ICC display differ from the ECS display. Due to limitations of the experimental display in simulating the ICC, color recommendations specific

to the ICC cannot always be made. These cases will be pointed out.

The Patriot display actually comprises three functional areas: the situation display, the alert message line, and the tabular display area. The display is illustrated in Figure 1. along with the control-indicator panel, keyboard, special function keys, and joystick. The situation display comprises the largest area on the display. It is a stylized map showing moving aircraft, or tracks, against a static background of objects such as search and track sectors, assets, safe passage corridors, weapons control volumes, and restricted and prohibited areas. For a complete listing of situation display objects and their respective line or shape coding, refer to Table 1. The alert message line is directly below the situation display and comprises a single line of text. It is where priority messages are displayed in succession. An example of a message on the alert message line is shown in Figure 2. The tabular display area is directly below the alert message line and comprises a rectangular area. It is where tables of information are displayed one at a time. One such table is the Engaged Data Table which contains the To-Be-Engaged Queue (TBEQ) and the Engaged Queue. The Engaged Data Table is illustrated in Figure 3.

Recommendations

The recommendations for each of these three display areas will now be described in turn.

Situation Display. The overall color scheme for the situation display is designed to reduce clutter and promote grouping of like objects. Areas and volumes will be drawn as filled polygons. Military color-meaning associations will be relied on to some extent: Objects associated with hostile information will be represented in reddish shades and those associated with friendly information will be represented in shades of blue. It should be made clear that when objects are represented in bluish shades, the blue is not a pure blue. Pure blue is not brought into focus as readily as other colors. Other objects will be represented in relatively neutral colors. Similar objects, that is, objects judged by Patriot operators to provide similar information, will be displayed in the same color, thus reducing the number of levels of color coding. In order to manage the potential color contrast problems, the background objects will be of relatively low intensity, that is, not very bright, and fairly unsaturated. Foreground objects will be of high intensity and highly saturated.

The current software configuration requires that the number and letter designators for each object be displayed in the same color as the object. This causes some contrast problems. In particular, a large area in a given color may contrast quite well with the its background; however, small letters and numbers in that same color are subject to the spreading effect (i.e., their color literally becomes diffused) and they do not contrast well. A fairly satisfactory resolution was achieved with the colors described in the remainder of this section. A more satisfactory resolution would be to display all letter and number designators in either black or white.

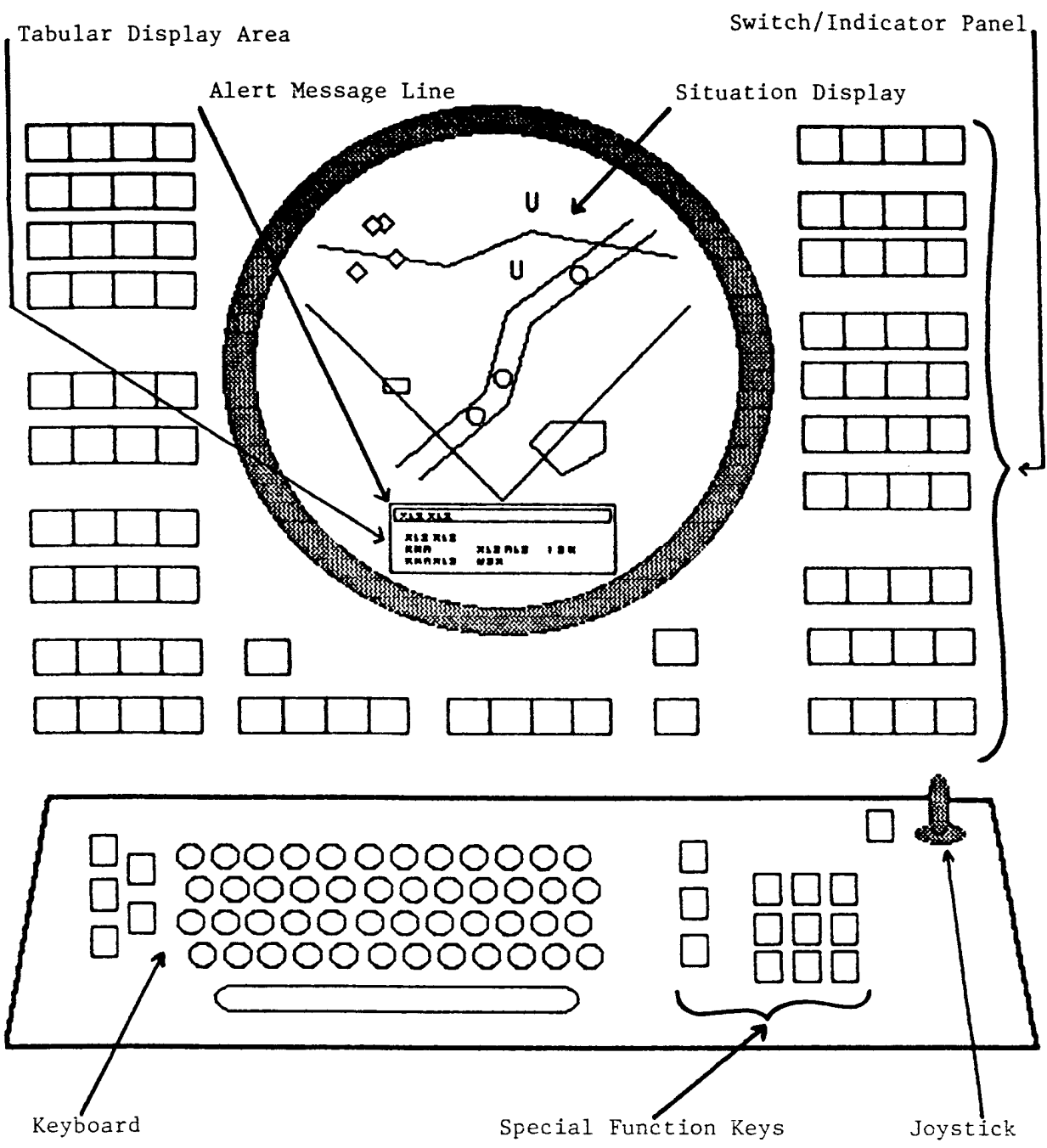


Figure 1. An illustration of the Patriot display showing the circular display comprising the situation display, the alert message line, and the tabular display area. Also shown are the switch/indicator panel, the keyboard, the special function keys, and the joystick. (TM-9-1430-600-10-1, 1983).

Table 1. Situation display objects and their shape or line code
(TM 9-1430-600-10-1, 1963).

Search Sector



Track Sector



Adjacent Sector Bounds



Weapons Control Volume-
Hold



Weapons Control Volume-
Tight



Weapons Control Volume-
Free



Restricted Volume



Prohibited Volume



Hostile Origin



Friendly Origin



Safe Passage Corridor



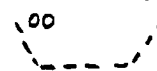
Defended Area/Point



FEBA



Masked Terrain



Range Rings



Table 1. continued. Situation display objects and their shape or line code
(TM 9-1430-600-10-1, 1963).

Fire Platoon Flag	P
Battalion Flag	B
Communications Relay Group	f
North Symbol	N
Friendly Track	O
Unknown Track	U
Hostile Track	◇
To-be-engaged Modifier	⬡
Engaged Modifier	⬢
Kill-assessed Modifier	#
Engage-hold Modifier	[H]
LNIP (Launch-Now Intercept-Point)	↘ 15
PIP (Predicted Intercept-Point)	↙ 15
Altitude/Threat Information	093
Track Number	001
Missile	0

Figure 2. Alert Message Line displaying the message that launcher station 1 is low on missiles and the word "more" to indicate that additional messages are waiting to be displayed when this one is cleared. (TM9-1430-600-10-1, 1983).

LS1 LOW MSL

MORE

Figure 3. The Engaged Data Tabular Display with entries on the TBEQ. No entries are shown on the Engaged Queue. (TGINO = target number; THRT = threat value; RT = release time; TLL = time to last launch; ENGSTAT = engaged status; ID/SZ = identification and size; MA = missile away; and TGO = time to go). (TM9-1430-600-10-1, 1983).

TGINO	THRT	RT	TLL	ENGSTAT	ID/SZ	_TGINO	MA	TGO	SZ
001	1	0	10		H				
002	1	15	45		H/2				
003	9	15	99		H				

1. Sector Bounds. The search and track sectors, currently designated by solid and dashed lines respectively, mark the coverage boundaries for a given fire platoon. For the ECS display, the two sets of lines can be eliminated and replaced with solidly colored areas by using the color-fill capability. This should reduce display clutter and make the coverage boundaries more readily visible at a glance. Since the search and track sectors functionally serve as background for all other display objects, the colors selected must be such that all other objects are visible when displayed on them. Two shades of bluish green or aqua were selected: a lighter shade for the search sector and a slightly darker shade for the track sector. No particular meanings are associated with aqua; therefore, it should provide a fairly neutral background.

Adjacent sector bounds are secondary information in relation to a fire platoon's own search and track sectors. Also, they are displayed only occasionally. For these reasons, they should continue to be displayed as dashed lines. A yellow-green, relatively neutral color, would be suitable, providing good visibility.

At the ICC, where search and track sectors for multiple fire platoons are displayed, using color fill quickly becomes unmanageable, therefore, color coded lines should be used.

2. Weapons Control Volumes. There are three types of weapons control volumes, hold, tight and free. All are displayed as polygons drawn with solid lines, differentiated by a number (1 - 9) and a letter designator (H, T or F). Although the three types of weapons control volumes convey different information, due to the number of other objects to be color coded, it is recommended that all three be displayed in a single color, still differentiated by the number and letter designator. Also, to reduce the number of similar appearing lines on the display, these volumes should be displayed as filled polygons. Regarding the color selection, while the volumes are displayed as foreground on the search and track sectors, they also serve as background to other objects. Therefore, the color should still be of relatively low intensity and unsaturated. A light orange, a reddish color, was selected for the experimental display.

3. Restricted and Prohibited Volumes. These two volumes are both used in the process of identifying unknown tracks as either friendly or hostile. Both are displayed as polygons drawn with dashed lines, differentiated by letter (R or P) and number (01 - 99) designators. The same reasoning applies to color use and selection here as with the weapons control volumes: Although there are two types of information, due to the number of other objects to be color coded, it is recommended that both the restricted and prohibited volumes be displayed in a single color, still differentiated by the letter and number designators. The volumes should be drawn as filled polygons with the result that they will be foreground on the search and track sectors and background for other objects. A pale pink, also a reddish color, was selected for the experimental display.

4. Origin Volumes. Hostile and friendly origin volumes are drawn as dashed polygons and are designated either "H" or "F" along with a two digit identifier (01-99). They too should be displayed as filled volumes, hostile

origins in a reddish shade and friendly origins in a bluish shade. On the experimental display, origin volumes are not displayed so no precise shades were chosen.

5. Safe Passage Corridor. Safe passage corridors are currently drawn on the display by two unconnected, parallel lines. Multiple corridors are differentiated by the letter "C" followed by two numbers (01 - 99). Clutter will be even further reduced by drawing the corridors as filled polygons. Since the corridors are usually narrower than weapons control volumes or prohibited or restricted volumes, corridors should be displayed as foreground over those volumes. Corridors will still be background to certain other objects. The color selected was a light blue, a color with a friendly connotation.

6. Defended Areas and Points. Defended areas and points are drawn as small squares or rectangles that may vary in size. A letter designator is used to distinguish them. Not so much to reduce clutter as to enhance visibility, they should be displayed as filled squares or rectangles which would be foreground to all the larger volumes and corridors. They would be background to just a few other objects, the most important of which are the tracks. A medium blue was selected for the experimental display.

7. FEBA (Forward Edge of the Battle Area). Different from the objects listed above, the FEBA is represented by a dashed line and is not a polygon. It provides important information and is always displayed. To make it highly visible over the other background colors, the FEBA should be drawn in a high intensity color. White was selected for the experimental display.

8. Masked Terrain. Areas of masked terrain are shown as short dashed lines and indicate terrain features which may block or mask air activity from the radar. Although they provide important information, they are not necessarily displayed continuously. Rather, they may be displayed only at critical times. Therefore, they should also be highly visible over the background. On the experimental display, bright orange, a reddish shade and an alerting color, was selected. Early results from the ARI Color Display Study indicate that this color does not contrast sufficiently with the background and needs to be made brighter.

9. Range Rings. Range rings are shown as dashed lines concentric to the fire platoon. Distances are shown by digits on the rightmost end of each ring. As with masked terrain, range rings are not usually displayed continuously. They too should be highly visible over the background colors. A dark purple or violet was selected. Again, as with the color selected for the masked terrain, early results of the study indicate that this color does not stand out sufficiently.

10. General Symbols. General symbols include the fire platoon flag, the battalion flag, the communications relay group flag and the North symbol. They are the smallest objects on the situation display. The first three symbols are judged by Patriot operators to provide related information and, therefore, should be displayed in the same color. Because of their small size, they need to be displayed in a very intense color to be visible. A yellow-green or chartreuse was selected for visibility even though it is more

a neutral color than one associated with the concept of friend.

The North symbol provides general geographic information unrelated to the other general symbols. Another difference is that it is displayed off to the side of the situation display. It should be displayed in an intense, neutral color. White was selected for the experimental display.

11. Tracks, Track Modifiers and Track Information. Tracks are shape coded according to identity, a "U" for unknown, a circle for friendly and a diamond for hostile. The track modifiers are shapes or letters written around or inside the tracks and are automatically displayed with the track. Supplemental track information is alphanumeric or shape coded information written alongside the tracks. It is displayed on request. (For a complete listing of the track modifiers and track information, refer to Table 1.) From this point on, all three kinds of information will be referred to simply as tracks. All three kinds of information for a given track are in tight physical proximity, and, of practical significance, all three kinds of information for a given track are required to be displayed in the same color by the current software.

The tracks provide the most important information on the situation display. Therefore, they need to be visible on all other background colors. For this reason and because of the small size of the symbols, the tracks should be displayed in colors of high intensity and high saturation. The colors should also be in agreement with military color-meaning associations. The colors chosen for the experimental display were yellow for unknown, dark blue for friend, and red for hostile. Early results of the ARI Color Display Study indicate that, while there are no serious problems with the selected colors, any steps to make them appear even more intense and more saturated would enhance the overall display.

Alert Message Line. The alert message line (see Figure 2.) is a single line of text below the situation display where important messages generated by the system or sent from the battalion to the fire platoon or vice versa are displayed. On the rightmost end of the line, the word "more" appears when there are messages waiting to be displayed when the current one is cleared. When this is the case, both the current message and the word "more" blink. A light gray was selected for the background and black was selected for the text. This color combination allows for sufficient contrast and yet is relatively easy on the eyes. Another reason for choosing these colors was so that the contrast was in the same direction as the situation display, dark, saturated foreground text on a light background. The experimental display does not provide for any color highlighting or color coding; however, it is possible that limited use of color to highlight the word "more" or to code high priority messages could be beneficial.

Tabular Display Area. Many different tables of information can be displayed in the tabular display area but since the experimental display allows for color coding beyond just background and text colors on only one table, the Engaged Data Table, only that one will be discussed in detail. Otherwise, the background and foreground text are the same colors as the alert message line, light gray and black, respectively.

The Engaged Data Table (see Figure 3.) contains information important for the engagement decision, especially the TBEQ. Based on tactical standard operating procedure, the decision was made to color code the entries on the TBEQ. When RT is equal to zero, the entry appears in red. When RT is greater than zero, but ATC is equal to one, the entry appears in yellow. All other entries appear in the standard text color, black. Early results on this coding are positive. The entries appearing in red seem to serve the alerting function well; however, entries virtually never appear in yellow. Future work might well examine another way to implement this level of the color code.

Summary Comments

To the observer, the most striking thing about the experimental Patriot color display is the way the areas and volumes are unified or integrated into solid objects. The display is no longer a maze of solid and dashed lines. The only remaining lines are objects which cannot be enclosed. Thus, color is serving its psychological function of grouping objects and reducing the clutter on the display. With respect to perceptual considerations, for the most part, the colors contrast well. And practical considerations have been addressed by keeping the number of colors to a minimum while having a sufficient number of colors to code the major categories of information. Of course, the real test of the experimental Patriot color display is whether it proves to be useful to the Patriot operator, the user. The study currently being conducted by ARI intends to answer that question by collecting data on the operators' opinions, their subjective assessments of workload, and operator/weapon system performance.

REFERENCES

TM 9-1430-600-10-1 Engagement Control Station, Guided Missile, Truck Mounted:
AN/MSQ-104 (October, 1983). Washington, D.C.: Headquarters Department
of the Army.

APPENDIX A
COLOR RESEARCH STUDY

The primary questions addressed by the ARI study are: (1) Does presenting the situation display, the alert message line and the supporting tabular displays in color provide an overall improvement in system performance by improving the operator's ability to understand and interact with the system? (2) And does the effectiveness of color vary with varying scenario difficulty? An outline of the key elements of the study follows:

1. Scope. The performance of operators in the Weapons Controller position, the position responsible for carrying out engagement decisions (i.e., firing on hostile tracks) at the Firing Battery Engagement Control Station (ECS), is being studied. The system is being operated in independent and autonomous modes. The test operators are given scenarios of varying levels of difficulty under the two display conditions where their task is always to defend the assets to the best of their abilities, to engage tracks in a timely fashion, and to conserve missile resources.

2. Test Operators. Trained Patriot operators, both 14E and 24T MOS's, are being used as test operators. Although the 24T is usually designated the Weapons Controller, the 14E is cross-trained as the Weapons Controller and, therefore, is suitably familiar with the task.

3. Equipment. A Raster Technologies Model One/80 is the primary piece of equipment being used for the study. It is used in conjunction with the color and pre-deployment build PTOS scenario software residing in the host computer, the Gould S.E.L. 32/77 under operating system 1.5B and 2.2. To complete the system, a keyboard and joystick are used as input devices as well as specially constructed button boxes used as the control-indicator panels.

4. Independent Variables. There are two independent variables of primary interest -- display type and scenario difficulty. These two variables are manipulated within subjects such that all test operators will use the two display types and will be presented with all levels of scenario difficulty. A third variable, MOS, is a between group comparison.

5. Dependent Measures. The dependent measures include objective and subjective measures. Individual cognitive and demographic data are being collected also.

a. Task Performance Measures (TPM's). These measures are based on keystroke data time-stamped at one second intervals. The length of time taken to perform certain critical tasks as identified through a task analysis will be computed from the start and stop times of the keystrokes. Typical task sequences will also be identified.

b. Summary Performance Measures (SPM's). These three measures, developed as a result of earlier ARI work (Hawley et al., 1982), are aggregate measures designed to capture overall system performance--man and machine--at the mission level. Mission level performance is measured in relation to general or long-term goals, to be contrasted with task level performance which is measured in relation to specific goals such as engaging an individual target. The SPM's are Defense of Assets, Air Space Defense, and Resource Conservation.

c. Subjective Workload Ratings. The Subjective Workload Assessment Technique (SWAT) (Reid, Shingledecker & Eggemeier, 1981) developed by the USAF Aerospace Medical Research Laboratory at Wright-Patterson Air Force Base is being used. It provides not only an overall assessment of subjective workload, but also an assessment of workload across three dimensions: time load, mental effort load, and psychological stress load. One further advantage of the technique is that individual differences in subjective assessment are taken into consideration in the scoring.

d. Scenario Difficulty Index (SDI) (Hawley et al., 1982). This index is different from objective scenario difficulty, which is manipulated as an independent variable. SDI is an average across a given scenario of the number of tracks to be dealt with weighted by the time pressure. SDI can be computed for each individual and it can also be computed without any operator intervention to provide a worst-case index against which to compare operator performance.

e. Subjective Ratings. Ratings include ratings of fatigue, glare, and flicker; display usability; and display preference. General comments also are being solicited.

f. Psychological Tests. A standard ARI battery of tests have been or will be administered. The battery includes tests of perceptual ability, cognitive ability, and personality.

g. Demographic Information. This information include age, rank, education/training level, recency of training, and training scores.

6. Procedure. Psychological tests are administered in large group sessions. Display and equipment familiarization and introduction to SWAT are accomplished in either individual or small group sessions of up to four test operators. It is important to ensure display and equipment familiarization since there are some configural differences between this system and actual Patriot equipment. Testing with the display is conducted in individual sessions. Each scenario will last about six minutes. Some of the subjective ratings and comments will be obtained after each scenario, some after each block, and some at the end of the entire testing session.

REFERENCES

Hawley, J. K., Brett, B. E., & Chapman, W. A. (1982). Optimizing operator performance on advanced training simulators: Specifications for operator performance assessment and workload quantification. Arlington, VA: US Army Research Institute Technical Report.

Reid, G. B., Shingledecker, C. A., & Eggemeier, F. T. (1981). Application of conjoint measurement to workload scale development. Proceedings of the Human Factors Society - 25th Annual Meeting, 522-526.

APPENDIX B
DESIGN CONSIDERATIONS

In this appendix general background information on perceptual and psychological aspects of color is presented. The criterion for inclusion here is the potential impact of the information on display design issues. Specific information from evaluations of a number of realistic color displays is presented also.

1. The three dimensions of color. People tend to perceive color unidimensionally although it actually varies along three dimensions—hue, brightness and saturation. Hue is what is generally thought of as color and refers to the wavelength of the light. Blue light is of relatively short wavelength; green and yellow are of medium wavelength; and red is of relatively long wavelength. Brightness refers to the intensity of the color. Saturation refers to the amount of white in the color. A highly saturated color has little or no white in it, for example, pure red. As a color becomes less saturated, it has more and more white in it, for example, pink.

2. The number of colors people can perceive. People can discriminate more than 150 different colors in side by side comparisons. However, the number of consistently identifiable colors in individual judgments only ranges up to about 15, with seven or eight being a reasonable upper limit in most cases (cf., Jones, 1962; McCallum & Rogers, 1982).

3. Color blindness. The numbers cited in the previous paragraph are based on people with normal color vision, but approximately 2% of the male population and about .03% of the female population can be classified as color blind, or more accurately, color deficient or having color confusions (Goldstein, 1984). Most of those have a red-green deficiency where they have difficulty discriminating between red and green. A very few have a blue-yellow deficiency where they have difficulty discriminating between blue and yellow. Only about 10 people out of a million can see no color at all, only shades of gray (Goldstein, 1984).

4. Differential focusing for different colors. Due to their different wavelengths, different colors require different degrees of accommodation or focusing. For most colors, the difference is negligible and not consciously noticeable; however, it is somewhat more evident with pure blue. Because of its very short wavelength, pure blue cannot be brought into focus simultaneously with other colors. In fact, it can be said to produce an almost myopic effect (Murch, 1984). Therefore, use of pure blue in an information display, in particular for small or fine-lined objects, is cautioned against.

5. Differential sensitivity to different colors. People are differentially sensitive to different wavelengths, with maximal sensitivity to yellow light (Goldstein, 1984). This means that less intensity is required for yellow to be perceived than for other colors, and that, at equal intensities, yellow will appear brighter than other colors.

6. Color contrast problems. The perception of a particular color is dependent not only on the color itself, but also on the color of the surrounding or background area. The background area can differ from a foreground object along one, two or all three color dimensions. There are four main types of color contrast problems, each of which will be discussed in turn.

a. Simultaneous color contrast. A hue difference can cause the phenomenon of simultaneous color contrast where the color of the object "shifts" away from the background color, towards its complement (Goldstein, 1984). For example, a purplish color will appear pinkish on a blue background and bluish on a pink background.

b. Spreading effect. A hue difference can also cause the spreading effect, (Goldstein, 1984; Walraven, 1984), conceptually the opposite effect of simultaneous color contrast: It is the tendency for an object to shift towards the surrounding or background color rather than towards its complement. With a color display, a spreading effect is created directly by the spreading and mixing of the light as it emanates from adjacent colors on the computer screen, an effect that is especially problematic with very small display objects.

c. Brightness contrast. A brightness contrast is where the color of the object appears brighter against a dim background or dimmer against a bright background.

d. Saturation contrast. Saturation contrast is where the color of the object appears more saturated against a whitish background and less saturated against a pure background (Spiker, Rogers & Cicinelli, 1983).

7. Color and attention. Color effectively extends the functional field of view, that "area around the central fixation point from which specific information can be extracted at a single glance" (Kraiss & Knaeuper, 1982). This occurs because color can be perceived automatically and effortlessly before attention is consciously focused on it (Kahneman, 1973; Triesman, 1982). Thus, preattentive color information can attract attention as in the case of a warning light. Or, in the case of a search task where a particular color has been designated the target, it can be used to direct attention.

8. Color and grouping. For the same reason that color is good for drawing and directing attention, color contributes to "good figure", that is, the perception of objects as unified wholes or the perceptual grouping of like objects. The preattentive color information is essentially added to the object or grouping information currently in focus so that a more integrated perception results. In other words, color helps to unify objects that are too large, or to group objects that are too widespread to be perceived in a single glance.

9. Color and clutter. Clutter refers to the number and kind of objects in the display. When a display is highly cluttered the usefulness of color for a search task is even greater than when the display is not cluttered, provided a reasonable number of colors is used. As the number of colors approaches either extreme--all objects one color or each object a different color--the usefulness of color as a search cue is lessened (Cahill & Carter, 1976; Carter & Cahill, 1979; Promisel, 1961).

10. Color and meaning. Colors do not have standard meanings like words. This is not to say that colors cannot have meanings, only that colors are not as explicitly identified with a particular meaning or interpretation as are words. Extended practice, however, will develop and strengthen color-meaning associations (Christ & Corso, 1983), for example, red for danger; yellow for caution; and green for go ahead (Smith & Aucella, 1983). In the military, typically, red is associated with hostile; yellow, amber or white with unknown; and blue or green with friendly.

11. Colors as an ordered set. Even though the color spectrum provides a natural ordering, colors are not ordinarily used as an ordered set like numbers or letters are. What is more, people report difficulty in using colors as an ordered set even when instructed to do so (Kanarick & Petersen, 1971). But, like color-meaning associations, color-order associations can be developed with extended practice.

12. Lessons learned from color used in other displays. To test the expected benefits of color in drawing, directing and focusing attention with a cluttered display, a number of realistic color displays have been implemented and evaluated. The main types of displays are battlefield maps, air traffic control displays, and jet cockpit displays. The primary lessons learned can be summarized in four main points which are discussed in the following sub-sections.

a. Search and identification as a function of how color is used. When color is added to a display, it can be added to the foreground objects, the background, or both. On the one hand, adding color coding to the foreground symbols on a battlefield map such as unit types or building sites has been shown to speed search for those symbols relative to using just the symbol coding. (Christner & Ray, 1961; Kafurke, 1981). And adding color coding can enhance the identification of differently colored, overlapping symbols such as on air traffic control displays or threat displays by helping to focus attention (Connolly, Spanier & Champion, 1975; Newman & Davis, 1962; Smith, 1963). On the other hand, adding color coding to the background, such as using color to shade altitude on a map, has been shown to reduce the time it takes to identify the altitude of an area, but to increase the time to find a specified x-y coordinate (Kempf & Poock, 1969). That is, when color was used to code the background rather than the foreground objects, color speeded identification of the background information, but slowed search for the foreground objects. It seems, then, that there is a tradeoff or balance between color coding the foreground objects and the background which must be considered in display design.

b. Grouping objects according to categories. Due to the role color plays in grouping like objects, color is useful for coding objects into categories, provided the code is at the level of the primary category. Two studies illustrate this point. In one study (Sidorsky, 1980) battlefield map symbols, each providing three levels of information--unit type, unit size, and unit readiness--were color coded according to one of the three levels. The task was to judge overall combat readiness. Color coding enhanced the judgments when the main category level, unit type, was color coded; however, it did not enhance performance when either of the two lower category levels, unit size and unit readiness, were color coded. In the other study (Patterson, Engleman, Najjar, & Corso, 1984), categories of aircraft altitude were color coded on an air traffic control display. The task was to identify aircraft flying too close in altitude. Color coding was not found to be an effective method for signaling possible collisions because endangered aircraft could be coded in the same color category or in different color categories. In both studies, then, when both between-category and within-category discriminations had to be made, color coding was not useful: Color is useful for coding categories only at the level of the primary category.

c. The effect on complex task performance. Most of the investigations of the effect of color on complex task performance have been conducted under the auspices of the US Air Force. Color has been implemented in simulated cockpits on terrain and airspace maps, and on navigation, weapons stores, system status, fuel level, and threat warning displays (Aretz, Calhoun, Kopala, Herron, & Reising, 1983; Kellogg, Kennedy, & Woodruff, 1984; Kopala et al., 1983; Stollings, 1984; Way, Hornsby, Gilmour, Edwards, & Hobbs, 1984). The types of tasks performed using these displays were flying; landing; carrying out bombing missions; checking the system status; responding to emergencies; locating, counting and comparing the number and kind of aircraft in a threat display; and avoiding and engaging enemy threat. Across studies it was found that color enhanced pilot performance of tasks such as checking system status information or counting and comparing the number and kind of aircraft; however, color only occasionally enhanced performance of flying, landing or bombing tasks. A possible explanation is that color benefits tasks where the response is determined by the human component (e.g., answering go or no go to a system status question) but not where the response is further determined by system features (e.g., landing an aircraft).

d. Reducing subjective workload. The subjective preference of the pilots in the studies cited in the previous paragraph was for color over monochrome or black and white. Of somewhat more interest is the fact that, generally, the pilots also felt that their performance was better with color. This suggests that color may reduce subjective workload, that is, how hard the task seems to be. Supporting evidence comes from a study (Luder & Barber, 1984) where the presence of color in a display enhanced performance on a secondary task both when the secondary task was performed concurrently with the primary task and when it was performed alone. The explanation given was that the mere awareness of the availability of the color information for the primary task served to free mental resources and reduce overall subjective workload. If this explanation is correct, it has important implications for display design.

REFERENCES

- Aretz, A. J., Calhoun, G. L., Kopala, C. J., Herron, E. L., & Reising, J. M. (1983). Computer generated color pictorial stores management displays for fighter aircraft (Report No. AFWAL-TR-83-3016). Wright-Patterson Air Force Base, OH: Flight Dynamics Laboratory, Air Force Wright Aeronautical Laboratories. (AD No. A131412)
- Cahill, M-C., & Carter, R. C. Jr. (1976). Color code size for searching displays of different densities. Human Factors, 18, 273-280.
- Carter, R. C., & Cahill, M-C. (1979). Regression models of search time for color-coded information displays. Human Factors, 21, 293-302.
- Christ, R. E., & Corso, G. M. (1983). The effects of extended practice on the evaluation of visual display codes. Human Factors, 25, 71-84.
- Christner, C. A., & Ray, H. W. (1961). An evaluation of the effect of selective combinations of target and background coding on map-reading performance - Experiment V. Human Factors, 3, 131-146.
- Connelly, D. W., Spanier, G., & Champion, F. (1975). Color display evaluation for air traffic control (Report No. FAA-RD-75-39). Washington, D.C.: US Department of Transportation FAA Systems Research & Development Service. (AD No. A010779)
- Goldstein, E. B. (1984). Sensation and perception (2nd Ed.). Belmont CA: Wadsworth Publishing Company.
- Jones, M. R. (1962). Color coding. Human Factors, 4, 355-365.
- Kafurke, P. (1981). An evaluation of the effectiveness of color coded tactical symbology applied to military maps. Monterey, CA: Naval Postgraduate School. (AD No. A109675)
- Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall Inc.
- Kanarick, A. F., & Petersen, R. C. (1971). Redundant color-coding and keeping track performance. Human Factors, 13, 183-188.
- Kellogg, R. S., Kennedy, R. S., & Woodruff, R. R. (1984). Comparison of color versus black-and-white as indicated by bombing and landing performance in the 2B35 TA-4J flight simulator (Report No. AFHRL-TR-84-22). Brooks Air Force Base, TX: Air Force Human Resources Laboratory. (AD No. A144674)
- Kempf, R. P., & Poock, G. K. (1969). Some effects of layer tinting of maps. Perceptual and Motor Skills, 29, 279-281.
- Kopala, C. J., Calhoun, G., Reising, J. M., & Herron, E. (1983). Symbology verification study (Report No. AFWAL-TR-82-3080). Wright-Patterson Air Force Base, OH: Flight Dynamics Laboratory (AFWAL/FIGR), Air Force Wright Aeronautical Laboratories (AFSC). (AD No. A131328)

- Kraiss, K. F. & Knaeuper, A. (1982). Using visual lobe area measurements to predict visual search performance. Human Factors, 24, 673-682.
- Luder C., & Barber, P. J. (1984). Redundant color coding on airborne CRT displays. Human Factors, 26, 19-32.
- McCallum, M. C. & Rogers, S. P. (1982). Application of coding methods in the development of symbology for a computer-generated topographic display for Army aviators. Fort Monmouth, NJ: Advanced Systems Division US Army Avionics R & D Activity. (AD No. A118142)
- Murch, G. M. (1984). Physiological principles for the effective use of color. IEEE Computer Graphics and Applications, 4, 49-54.
- Newman, K. M., & Davis, A. R. (1962). Non-redundant color, brightness, and flashing rate encoding of geometric symbols on a visual display. Journal of Engineering Psychology, 1, 47-67.
- Patterson, M. J., Engleman, W. J., Najjar, L. J., & Corso, G. M. (1984). The effects of redundancy, target density and practice of a complex visual search task. Proceedings of the Human Factors Society - 28th Annual Meeting, 299-302.
- Promisel, D. M. (1961). Visual target location as a function of number and kind of competing signals. Journal of Applied Psychology, 45, 420-427.
- Sidorsky, R. C. (1980). Color coding in tactical displays: Help or hindrance? Arlington, VA: US Army Research Institute Technical Paper in Preparation.
- Smith, S. L., (1963). Legibility of overprinted symbols in multi-colored displays. Journal of Engineering Psychology, 2, 82-96.
- Smith, S. L., & Aucella, A. F. (1983). Design guidelines for the user interface to computer-based information systems (Report No. ESD-TR-83-122). Bedford, MA: The MITRE Corporation.
- Spiker, V. A., Rogers, S. P., & Cicinelli, J. (1983). Identification of colored stimuli on a computer-generated raster display as a function of luminance, foreground color, background color, and stimulus size (Report No. 459-6). Santa Barbara, CA: Anacapa Sciences, Inc.
- Stollings, M. N. (1984). Information processing load of graphic versus alphanumeric weapon format displays for advanced fighter cockpits (Report No. TR-84-3037). Wright-Patterson Air Force Base OH: Flight Dynamics Laboratory, Air Force Wright Aeronautical Laboratories.
- Triesman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. Journal of Experimental Psychology: Human Perception and Performance, 8, 194-214.

OPERATOR WORKLOAD (OWL) ASSESSMENT PROGRAM FOR THE ARMY

VALIDATION AND ANALYSIS PLANS FOR THREE SYSTEMS

[ATHS, Aquila, LOS-F(H)]

A. C. Bittner, Jr.
A. Zaklad
A. O. Dick
R. J. Wherry, Jr.
E. D. Herman
J. Bulger
P. M. Linton
R. J. Lysaght
T. W. Dennison



**U.S. Army Research Institute
for the Behavioral and Social Sciences**
5001 Eisenhower Avenue, Alexandria VA 22333

This working paper is an unofficial document intended for limited distribution to obtain comments. The views, opinions, and/or findings contained in this document are those of the author(s) and should not be construed as the official position of ARI or as an official Department of the Army position, policy, or decision, unless so designated by other official documentation.

Draft Technical Report 2075-3B

OPERATOR WORKLOAD (OWL)
ASSESSMENT PROGRAM FOR THE ARMY:
VALIDATION AND ANALYSIS PLANS
FOR THREE SYSTEMS
[ATHS, AQUILA, LOS-F(H)]

4 September 1987

Submitted to:

Dr. Richard Christ
Army Research Institute
P.O. Box 6507
Ft. Bliss, TX 79906-0057

Prepared by:

Alvah C. Bittner, Jr., Ph.D.	John Bulger, MS, MBA
Allen Zaklad, Ph.D.	Paul M. Linton, BSEE
A.O. Dick, Ph.D.	Robert J. Lysaght, Ph.D.
Robert J. Wherry, Jr., Ph.D.	T.W. Dennison, MA
Eric D. Herman	

Task Report
Contract No. MDA903-86-C-0384
Item No. 9002AJ



ABSTRACT

The purpose of this report is to layout plans for the validation of operator workload (OWL) measures and assessment techniques on three separate Army systems. It is a companion to the volume where these measures are reviewed and discussed in detail, *Comprehensive Review and Evaluation of Operator Workload Methodologies* (Lysaght et al., 1987).

Three systems in different stages of the materiel life cycle are described and OWL validation plans are developed. The approach for each system differs from the others where necessary and is similar where possible. Each validation plan represents an attempt to provide double duty: Validation is the primary goal but providing useful information to the system developers is also a primary goal. To facilitate the application of the results, we have met with various individuals concerning each system, and will continue contact with them, to understand fully their respective needs. These meetings, and our review of OWL measures have lead to the development of objectives for each system and implementation plans consonant with the development schedules.



ACKNOWLEDGEMENTS

This project has received leadership, guidance, inspiration and support from Richard E. Christ and his colleagues at the Army Research Institute (ARI).

Special thanks are due to Mr. Frank Brown, Horizens Technology Inc; Mr. Edwin Smootz, ARI - Fort Hood, Texas; the officers and men of the Target Acquisition Department, US Army Field Artillery School and, D Battery (RPV), 2/2d Field Artillery Battalion; and Mr. Tom Metzler, ARI - AVSCOM, for their help in developing our understanding of both the systems to be studied and the development programs associated with those systems. We also recognize all the people in the AMC project offices and TRADOC organizations associated with the systems to be studied who took the time from their busy schedules to provide us with cooperation and information. This report could not have been prepared without their assistance.

And, finally, many thanks to our colleagues at Analytics for their cooperation, especially Pam Lawfer and Heidi Brauer who assisted in proofreading and formatting the text.

The views, opinions, and/or findings contained in this report are those of the authors and should not be construed as an Official Department of the Army position, policy or decision.



TABLE OF CONTENTS

1.	INTRODUCTION.....	1
1.1	Program Background.....	1
1.2	Selection of Systems and Assessments	4
1.3	Organization of Report	6
2.	METHODOLOGY.....	7
2.1	General Approach.....	7
2.2	Analysis of Sensitivities.....	8
2.2.1	Sensitivity, Cost, and Efficiency.....	8
2.2.1.1	Sensitivity	8
2.2.1.2	Costs and Efficiency	9
2.2.1.3	Correlational Efficiencies.....	10
2.2.1.4	Statistical Comparison of Correlations.....	10
2.2.2	System Condition Variations	11
2.2.3	Individual Differences.....	11
2.2.4	Undesirable Sensitivities.....	12
2.3	Analysis of Assumptions.....	13
3.	AIRBORNE TARGET HANDOVER SYSTEM.....	15
3.1	ATHS/AI System Overview.....	15
3.1.1	Description.....	17
3.1.2	ATHS/AI in Mission Context	22
3.1.3	Previous Efforts Related to OWL.....	23
3.2	Program Schedule	24
3.3	Validation Plan	27
3.4	Summary.....	30

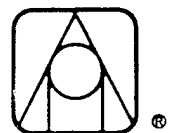


TABLE OF CONTENTS (CONTINUED)

4. AQUILA REMOTELY PILOTED VEHICLE (RPV) TARGET ACQUISITION / DESIGNATION AND AERIAL RECONNAISSANCE SYSTEM (TADARS) 31

4.1 Aquila System Overview..... 31

4.1.1 Aquila Mission Roles..... 31

4.1.2 Aquila System Description..... 32

4.1.3 Aquila Field Deployment..... 34

4.1.4 Conduct of Aquila Missions..... 36

4.1.5 Functional Description of GCS Displays and Controls 37

4.1.6 Previous Efforts Related to OWL..... 41

4.2 Program Schedule 42

4.3 Validation plan for Aquila..... 44

4.3.1 Objectives of Validation Plan..... 44

5. LINE OF SIGHT-FORWARD (HEAVY)..... 48

5.1 LOS-F(H) System Overview..... 48

5.1.1 FAADS Program Overview 49

5.1.2 LOS-F(H) System and Mission..... 51

5.1.3 Functional Description of the LOS-F(H) Display-Control Layout..... 55

5.1.4 Current and Previous Efforts Related to OWL 59

5.2 Program Schedule 61

5.3 Validation plan for LOS-F(H)..... 64

5.3.1 Objectives of Validation Plan..... 64

5.3.2 Experimental Design for LOS-F(H)..... 67

5.3.2.1 Subjects..... 68

5.3.2.2 Missions 68

5.3.2.3 Workload Measures 69

5.3.2.4 Data and Analysis..... 69

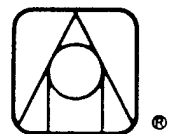


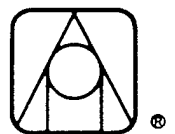
TABLE OF CONTENTS (CONTINUED)

6. CONCLUSIONS 71

 6.1 OWL Analysis..... 71

 6.2 Management Plan 71

7. REFERENCES 74



1. INTRODUCTION

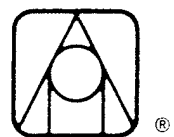
This report is one of a series describing a program for the development and validation of a methodology for estimation and evaluation of operator workload (OWL) in Army systems. It presents the results of Subtask 3.2 of Analytics' contract with the Army Research Institute (ARI). Subtask 3.1, a companion effort concerned with a comprehensive evaluation of workload measures, is documented in a separate report (Lysaght, Hill, Dick, Plamondon, Linton, Wierwille, Zaklad & Bittner, 1987). The present subtask is concerned with delineating:

- OWL measures that are recommended for estimation and evaluation of the three systems selected earlier by the Army -- the Aquila Remotely Piloted Vehicle (RPV), Airborne Target Handover System (ATHS), and the Line of Sight-Forward (Heavy) (LOS-F(H)) air defense system.
- Plans for performing OWL estimation and evaluation within the context of the ongoing programs for Aquila, ATHS, and LOS-F(H).

Delineation of these measures provides the framework for later efforts. Task 4 will involve the implementation of the planned estimations, evaluations and validations for Aquila, ATHS and LOS-F(H). In turn, these analyses and validation efforts will provide illustrations of applications and augment the description of our methods in the OWL Handbooks to be prepared during Task 5. The broad purpose of this report is to delineate plans for OWL analysis and validation studies of the ATHS, Aquila, and LOS-F(H).

1.1 Program Background

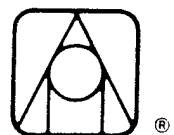
The validation and analysis plans developed in this report build upon programmatic evaluation of Army requirements and user needs regarding OWL (Bittner et al., 1987; Hill, Lysaght, et al., 1987; Hill, Plamondon et al., 1987). The results of this evaluation have directed our plans (1) toward filling significant voids in the current technology base for OWL, and (2) for the assessment of the selected systems (i.e., ATHS, Aquila and LOS-



F(H)). These plans will be described to provide the context for the more extensive and specific discussions of methodology and plans in later sections.

Hill and colleagues observed significant gaps in the current technological base for both the analytical and empirical methods, outlined in Figure 1.1-1 (Hill, Lysaght et al., 1987; Hill, Plamondon et al., 1987). They noted that the operational validities of the analytic methods had not been well established. An exception is Pro-SWAT, an expert opinion method which has been partially validated for some flight-oriented applications (Eggleston, 1984; Kuperman, 1985; Reid et al., 1984). This lack of systematic validation is unfortunate because in surveying Army system developers, it was found that there is a strong need for low cost techniques for input of workload issues into early design documents (Hill, Lysaght et al., 1987; Hill, Plamondon et al., 1987). As pointed out, costs of design changes increase rapidly as development continues. After a physical system has been implemented, there are frequently orders of magnitude greater than that during early development. The importance of the analytical methods for early design was the motivation for their recent characterization by Hill, Lysaght et al., (1987). The high potential of low-resource analytical techniques provided the motivation for focus on this class of techniques in the plans for Aquila, ATHS and LOS-F(H).

Hill and colleagues also observed that the current OWL analytical and empirical literature does not systematically address individual differences as they are of concern to the Army. Until recently, individual differences in subjective OWL data appear to have been largely attributed to differences in personal definitions or responses to aspects of workload (e.g., Hart, Childress & Hauser, 1982; Reid, Eggemeier & Nygren, 1982). More recently, attention has begun to focus on individual differences reflecting personality or cognitive-style variables (e.g., Damos, 1985), but has still lagged in recognizing differences related to information-processing (Damos, 1984). However, Personnel, one area of the Army Manpower and Personnel Integration (MANPRINT) initiative, is particularly concerned with information-processing related differences (e.g., Armed Services Vocational Aptitude Battery [ASVAB]). This interest is focused on their relationship to the operation, maintenance and support of systems (Hill et al., 1987, 2-35ff). ASVAB and related individual differences will be evaluated to the extent possible in the later presented plans. Individual differences in processing capabilities, it is noteworthy, have also not been adequately treated by the OWL analytical models, although addressed conceptually by some developers (e.g., Lane, Strieb & Leyland, 1979). However, the MANPRINT Method Program is pointing toward systematic treatments of individual



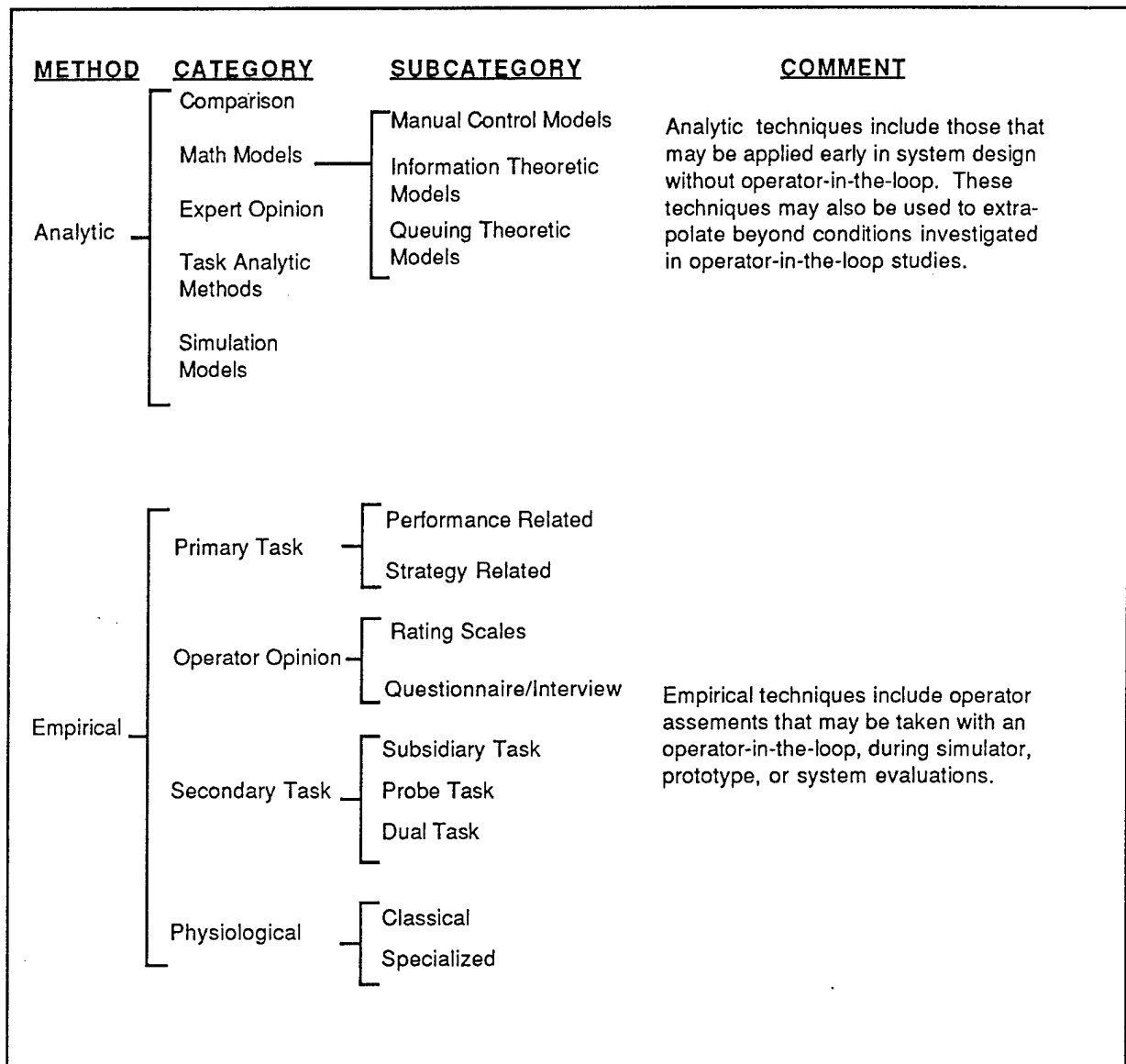
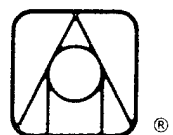


Figure 1.1-1. Taxonomy of Workload Assessment Techniques

differences in some OWL analytic technique models, such as MicroSAINT and HOS (e.g., Glenn, Dick & Bittner, 1987). Clearly, OWL information-processing related individual differences are an area of substantial interest to the Army. The importance of individual differences provides the motivation for their emphasis in the plans for Aquila, ATHS and LOS-F(H).



1.2 Selection of Systems and Assessments

The selection of the systems for validation and analysis were built upon evaluation of requirements and user needs regarding OWL (Bittner et al., 1987; Hill Lysaght et al., 1987). To ensure evaluation of a spectrum representative of combat systems, candidates were identified and evaluated in collaboration with cognizant system representatives and then selected to represent a range of diverse workload features. Table 1.2-1 summarizes estimates of salient features of operator workload developed during the selection of systems (Hill, Lysaght et al., 1987, 5.1ff.). Additionally, it may be noted that during this earlier evaluation "physical workload" was also evaluated and that, with one exception, estimates were equally low across system operators. The one exception was the LOS-F(H) Gunner where the level was estimated as potentially moderate during missile loading operations. Estimated perceptual, cognitive, motor and communication workload levels may altogether be seen to span a wide range.

Validation plans were prepared individually tailored to each selected system. In particular, plans were formulated to include OWL measures that offered the greatest potential utility and impact for the specific combat systems selected. For each selected system, LOS-F(H), ATHS and Aquila, a goal was to utilize batteries of both analytic and empirical techniques. In particular, a family of relatively low-cost expert opinion techniques (Pro-SWAT, Pro-TLX, etc.) were included in order to assess their relative utility in providing similar types of information before operator-in-the-loop simulations. Where possible, these were augmented to include other analytic methods such as Comparison, Task-Analytic Methods, or Simulation. A parallel family of low-cost, empirical subjective techniques (e.g., SWAT, TLX, etc.) were also included in order to assess their relative utility in providing information during operator-in-the-loop evaluations. The validation plans were altogether structured to; (1) ensure benefit for the selected systems; (2) be reflective of crew workload issues to the extent possible within the system evaluation opportunities; and (3) provide valuable gap-filling information on OWL methods. Validation plans presented in this report will be implemented with more detailed data and analysis as this data becomes available.

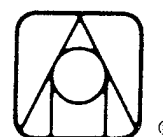
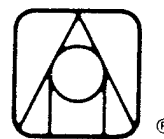


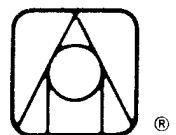
Table 1.2-1. OWL Overview of the Selected System

SYSTEM	OPERATOR	ENVIRON IMPACT	OWL CATEGORY LEVEL			
			PERCEPTUAL	COGNITIVE	MOTOR	COMMUNICATION
LOS-F(H)	GUNNER	MOD	MOD - HIGH	MOD	MOD - HIGH	LOW - MOD
	SQUAD LEADER	MOD	MOD - HIGH	MOD - HIGH	LOW	LOW - MOD
ATHS	COPILOT/ GUNNER	MOD - HIGH	LOW - MOD ?	MOD - HIGH	MOD - HIGH	LOW ?
AQUILA	VEHICLE OPERATOR	LOW	MOD - HIGH ?	MOD - HIGH ?	HIGH ?	LOW - MOD ?
	MISSION COMMANDER	LOW	MOD - HIGH ?	MOD - HIGH ?	MOD ?	MOD - HIGH ?
	PAYLOAD OPERATOR	LOW	MOD - HIGH ?	MOD - HIGH ?	HIGH ?	LOW - MOD ?



1.3 Organization of Report

The body of this report presents validation and analysis plans prepared for each of three systems as part of Task 3 (i.e., ATHS, Aquila and LOS-F(H)). Specifically addressed are: (1) the measures that are recommended for estimating and evaluating OWL for each of these systems as well as; (2) plans for conducting analysis within the context of their ongoing programs. To set the stage for discussions of the plans for each system, Section 2 considers general methodological issues including the assessment of measure sensitivity and efficiency and the testing of assumptions which underlie the data analysis plan. Subsequently, plans for the ATHS, Aquila and LOS-F(H) are developed in Sections 3, 4 and 5, respectively. Each of these sections initially overviews respective systems, mission contexts and previous efforts related to OWL. Finally in each of these sections, validation plans are developed in light of their system program schedules. Section 6 concludes with a synoptic discussion of the validation efforts and our overall strategy for their implementation.



2 . METHODOLOGY

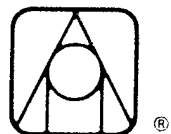
The methodological framework for the later validation efforts is delineated in this section. In Subsection 2.1, the context and general approach for the evaluation planning is characterized. Subsection 2.2 considers the evaluation of technique sensitivities to system condition variations, individual differences, as well as other "undesirable" variables. Subsection 2.3 discusses the evaluation of the assumptions which underlie the statistical procedures typically used in system evaluations of OWL.

2.1 General Approach

The development of the selected system evaluation plans has and continues to be a dynamically evolving process. Since our initial system selections were made (Hill, Lysaght et al., 1987), program schedules have changed with unfolding system developments. These changes are typical of combat system developments and offer realistically challenging environments for our evaluations of OWL. Our approach has been, while following these developments, to; (1) become intimately familiar with salient system characteristics and; (2) identify opportunities for significant benefits to the selected systems. In the following discussions of each of the selected systems, the results of this approach are presented, in terms of:

- System Descriptions;
- Outlines of previous workload-related studies;
- Current program schedules;
- OWL Evaluation schedules; and
- ARI OWL Objectives.

A system-specific experimental plan is provided which has been developed to begin to meet the objectives for the system with the most stable schedule [viz., LOS-F(H)]. Such specific plans will be reviewed by cognizant system personnel before implementation. It is



within the context of such specific plans that the evaluations of the below discussed sensitivities and assumptions will later be conducted.

2.2 Analysis of Sensitivities

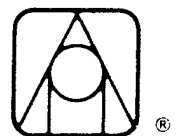
This section considers the evaluation of technique sensitivities. Subsection 2.2.1 discusses efficiency as well as provides other general background regarding sensitivity and its assessment. Subsequently, Subsections 2.2.2 and 2.2.3 focus respectively on desired system condition variations and individual differences sensitivities. Subsection 2.2.4 concludes with considerations of sensitivities to undesirable variables.

2.2.1 Sensitivity, Cost, and Efficiency

This subsection discusses a conceptual framework from which to view the analysis of sensitivities of the various measures of OWL. This framework, although not exhaustive, does broadly delineate the nature of specific data analysis procedures which we plan to apply across systems. Subsequent discussions of sensitivities to conditions, individual differences and other variables assume some familiarity with the issues raised in this subsection. For the general reader, the key points concern the importance of and the interrelationships among technique sensitivity, cost, and efficiency. The discussion tersely cites the important results and sources for the reader interested in the technical details

2.2.1.1 Sensitivity

The *sensitivity* of a measure may be defined in terms of the "value or amount of relevant information it provides." Statistically, sensitivity may be approached from a number of *almost* isomorphic statistical approaches. In the OWL literature, for example, the sensitivity of measures has been approached via Analysis-Of-Variance (ANOVA) in a number of reported efforts (e.g., Wierwille, Casali, Conner, & Rahimi, 1985). Under this approach, two or more measures are qualitatively compared in terms of statistical significances under common experimental conditions with varied "workloads". In particular, a measure is judged to be more sensitive than a second, if that measure demonstrates significance and the second does not. Likewise they are judged roughly comparable, if both or neither is significant. This approach, it may be noted, is not



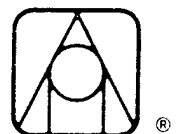
definitive because statistical sampling variations in two equally sensitive measures may result in one demonstrating significance and the second not.

Multifactor ANOVA has been applied to provide a more direct quantitative comparison of measures. With measures (M) representing levels of one factor and workload conditions (W) levels of a second, the significance (and follow-up analysis) of the (MxW) interaction in principle provides a direct comparison of the sensitivity of the measures. A significant MxW interaction is taken to indicate differential sensitivity to the workload conditions (and non-significance reflects effective equivalence). Unfortunately, this approach generally requires both that the measures be statistically commensurable, and that certain statistical adjustments must be made. Such adjustments are not common in the workload literature (cf., Huynh & Feldt). More importantly, this approach does not consider the differential costs and efficiencies associated with the various workload measures.

2.2.1.2 Costs and Efficiency

Costs are always associated with collection of data. The kind of cost of particular interest in the following discussions is the *amount of time needed to acquire information*. Other kinds of costs may be identified including those involved with, for example, the obtaining and use of special equipment and supporting technicians. These non-temporal costs are important but are not of interest in the current discussion. Data acquisition time is of interest because it generally may be analytically traded-off for both measure reliability and consequent sensitivity. For most types of continuous data (e.g., physiological), the component which may be traded-off is the time required for sampling units of data (cf., Bittner et al., 1986). For discrete methods such as rating scales, the component of particular interest for reliability trade-offs is the time required to produce a specific unit sample of data.

Trade-offs of temporal cost and reliability are important to consider. A given workload measure, for example, may provide slightly more "valid" data than another measure, but at a far greater acquisition cost. To evaluate this type of trade-off, it is necessary to consider the sensitivity of a measure relative to its temporal costs, i.e., *efficiency*. One of the objectives of the application studies will be to determine and compare the relative efficiencies of various measures of OWL using correlational approaches.



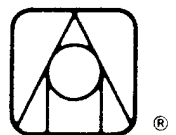
2.2.1.3 Correlational Efficiencies

Correlations that may be adjusted for temporal costs can be estimated by both ANOVA and averaged-correlation-based approaches. Illustrating both of these approaches, Winer (1962, pp. 129-130) estimates the reliability of judges from a table of repeated measures and contrasts his results with one form of averaged-correlation (bottom of p. 130). Although not shown by Winer, analogous correlations reflecting sensitivity to conditions may also be computed from such a table after rotating the roles of judges and conditions. This ANOVA-based approach, it is noteworthy, is conceptually similar to one more comprehensively described elsewhere (Chronbach, Gleser, Nada & Rajaratnam, 1972). It is also noteworthy that an approach to averaging correlations has been considered elsewhere (Dunlap et al., 1986). Correlational sensitivities may be estimated by a number of methods that can provide a basis for efficiency adjustments.

The correlations reflecting sensitivity may be further (analytically) adjusted for acquisition times based upon the classical relationship of Spearman-Brown (cf., Allen & Yen, 1979, pp. 85-88; Winer, 1962, p. 127; 1971, p. 286). This adjustment has been applied for comparative evaluations of time-adjusted "reliability-efficiencies" of more than 114 human capability measures (Bittner et al., 1986, p. 687ff.). Mathematically combining reliability-efficiencies with the correction-for-attenuation equation (Allen & Yen, pp. 98-99), and time-adjusted condition "sensitivity-efficiencies" may also be evaluated. Interestingly, the interplay between reliability and sensitivity implied by this adjustment has been recently addressed by Lane and associates in the context of the failures of operational performance assessments (cf., Lysaght et al., 1987). Sensitivity- and reliability-efficiencies are metrics of particular interest as part of our later system application efforts.

2.2.1.4 Statistical Comparison of Correlations.

Tests for comparison of simple independent bivariate-normal correlations are well known and straightforward (e.g., Anderson, 1984, pp. 122-124). When non-independent and non-normally distributed, the comparison of correlations is more problematic. For adjusted correlations (e.g., those described above) analytic corrections for standard results may be developed; however, non-independence and non-normal distributions make such developments extremely problematical. In light of these potential problems, we have selected a general approach for comparison of correlations which appears ideal for the breadth of potential problems: Jackknife Methods (cf., Hinkley, 1983). This class of



methods has been previously applied to similar problematic variance comparisons in multifactor repeated measures experiments (Carter & Bittner, 1982). More importantly, this approach has been previously applied for comparison of non-independent, non-normal, and adjusted correlations. Jackknife Methods are currently envisioned as a tool for potential use in evaluating system condition variation and individual differences sensitivities.

2.2.2 System Condition Variations

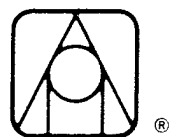
OWL measures should have *sensitivities* to a broad variety of operator workload condition variations in the context of issues implied by MANPRINT (Hill, Lysaght et al., 1987). This variety would include:

- *Within Mission Scenarios* -- variations experienced with changing mission segments,
- *Between Systems* -- variations arising from differing generations of a common system or divergent alternatives designed to meet the same requirement,
- *With Differing Levels of Training and Practice* -- qualitative differences in training as well as quantitative difference in its amount (these could have been also treated as "individual differences"), and
- *Combinations of Interest* -- selected combinations which could be usefully traded-off against each other (e.g., Equipment vs. Training).

Not all of these have been addressed in previous comparisons of methods although it may be argued that differential results are possible. In the following sections describing specific system applications, some combinations of common interest for combat systems will be evaluated based upon available opportunities and resources. System condition variation sensitivities and sensitivity efficiencies will be generally addressed in Subsection 2.2.1.

2.2.3 Individual Differences

The relationship between operator workload measures and individual differences has not been adequately addressed in previous evaluation efforts (cf., Section 1.1). To provide a framework for remedying this, a general assessment of various sensitivities to non-specific differences has been presented (Subsection 2.2.1). The discussed reliability correlations and efficiencies reflect the magnitudes of differences, but they do not indicate



their nature or sources. The identification of the nature of such differences has utility for selection and assignment purposes, particularly when separated into components that are:

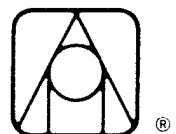
- *Information-processing related* -- these would include ASVAB and related Military Occupational Specialty (MOS) selection variables of interest to the Army, and
- *Other than information-processing related* --- these would include those that cannot be related to such differences.

In our analyses of individual differences, it is our intention to separate individual differences into these two components to the extent possible. Conceptually, stepwise multiple-correlation appears particularly suitable for identifying the proportion-of-variance of individual differences which may be attributed to the scores from the ASVAB (viz. R^2). Likewise, the complementary proportion which may not be attributed to such differences may also be computed ($1-R^2$). This approach, however, is not generally practicable for the relatively meager numbers of subjects expected for most of our system applications. A more directed method will usually be employed which sorts the individual differences into two orthogonal components: MOS-Selection-Composite Related and Other Related. This MOS composite approach may generally be applied to as few as three operators of a common MOS.

2.2.4 Undesirable Sensitivities

Workload estimation techniques should ideally be sensitive to system condition and individual differences, but insensitive to unrelated or extraneous factors. Unfortunately, extraneous factors are well-known to impact the results of subjective estimations such as ratings. These are considered here because of their potential impacts on our later system evaluations as well as their implications regarding the validities of the measures. Four undesirable factors are delineated in the following:

- *Order of ratings* -- this effect arises from various causes related to the sequence in which ratings are made. For example, items presented early must be decided upon without the "context" available for the subsequent ratings. Ratings of items may also "carry-over" into subsequent ratings.
- *Experimenter bias*-- these are the results of influence or expectancies brought to bear by the person who instructs, monitors, or demands the ratings to be completed. Things stressed by such persons can significantly influence the outcomes of the ratings to favor those desired by the experimenter. Reactions of experimenters -- both conscious and



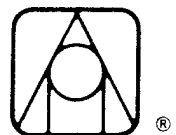
unconscious -- can also provide cues to raters that unduly influence the raters.

- *Conscious rater bias*-- this is the result of a rater attempting to provide misinformation or to exaggerate claims so as to influence the outcome of decisions for which the rating information is being collected. For example, if a rater wanted to have the first of two alternatives selected, the rater might rate that alternative more favorably than it objectively deserves.
- *Subconscious rater bias* --this is similar to conscious bias, but the rater is unaware that his ratings are biased. A particular form of subconscious bias is called *halo*. Positive-halo is the tendency to report that all facets of something relatively favorably because some few facets of it are attractive to the rater. Negative halo is the reverse tendency to report that all facets of something relatively unfavorably because of some feature unattractive to the rater. Halo in ratings tend to make all rated facets appear to be more interrelated than they actually would be if ratings were unbiased.

The impact of these effects may be controlled (or at least evaluated) in a number of ways. In the case of **order effects**, sequence and carry-over effects may be evaluated and to some extent controlled through the use of "balance for residual effects designs." This is the approach which we intend to take during the three system applications, where practicable. **Experimenter bias** may be somewhat controlled by care taken to ensure that the rating administrator is objective about what is being rated. The magnitude of these effects may also be explored by use and evaluation of designs in which administrators are systematically balanced. Objectivity is a goal which hopefully will be maintained by our evaluation teams. However, we are also exploring opportunities to evaluate experimenter effects through controlled variations. **Conscious and subconscious bias** are problematical. However, rater awareness via instructions of the nature of conscious and subconscious bias may partially alleviate this problem. For evaluation of these bias effects, systematic selections of variations in instructions and populations of raters may offer some means. Both of the approaches have been explored, but presently only instructional control appears implementable. Altogether, our approach toward these undesirable sensitivities represents a mix of control and evaluation strategies.

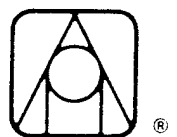
2.3 Analysis of Assumptions

Statistical tests will frequently be applied during our system evaluation efforts to identify significant effects related to workload variations and individual differences. Mathematically, the sampling distributions of these tests are based upon assumptions which



are critical to the validity of significance levels based on these distributions. Fortunately, many classes of tests-of-significance are robust to many assumption violations (e.g., Scheffe, 1959) and we have attempted to take advantage of these in our approach (e.g., Subsection 2.2.1.4).

Repeated measures designs often present particular problems because of violations of strong assumptions regarding subject consistencies across repetitions (i.e., compound symmetry). These have been shown to be substantially violated in the majority of 114 human performance measures (Bittner et al., 1986). Interestingly, our review of several sets of OWL rating data also appeared to reveal substantial violations of compound symmetry. However, visual inspections during the Wierwille et al. (1985) effort failed to detect such gross violations (Wierwille, 1987, personal communication). This divergence of results has prompted our plans to both evaluate symmetry assumptions during our system evaluation and adjust our results by appropriate corrections where practicable (Huynh & Feldt, 1976; 1980).



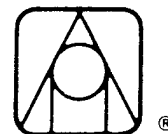
3. AIRBORNE TARGET HANDOVER SYSTEM

The Airborne Target Handover System (ATHS) is discussed in this section in the context of the validation efforts. The discussion has three parts. In Section 3.1, the system is described in the context of helicopter missions and previous analyses relevant to OWL. In Section 3.2, the program schedule is presented together with the schedule for validation plans. Finally, the validation plans are presented in the context of three program objectives in Section 3.3.

3.1 ATHS/AI System Overview

The Airborne Target Handover System (ATHS) Avionics Integration (AI) is a battlefield mission computer and a digital communications network in the most general form. ATHS/AI is a separate but integrated system in multifunction hardware. It is integrated with the Command, Navigation and Identification (CNI) subsystems via the 1553 bus as shown in Figure 3.1-1. It is accessed through the use of the generic Control and Display Unit (CDU) as are other functions. The purpose of ATHS/AI is to provide "reliable, rapid, and Electronic Countermeasure resistant target and battle management information transfer between similarly equipped scout helicopters, attack helicopters, field artillery systems and command centers (ATHS/AI System Specification, Attachment 11, RFP DAAJ09-87-R-A215, p. 9). This capability enables command and firing element crews to exchange target and other mission essential information using a short data-burst rather than voice communications. This data-burst minimizes the possibility of jamming, lessens the probability of detection, and speeds the transfer of accurate battle information.

The system elements may be installed on ground based or airborne vehicles to provide digital communications networking capability for varied combat elements. These include close air support, artillery, infantry, armor, and air defense. Thus, selection of this particular system for detailed workload analysis will have wide interest and applicability



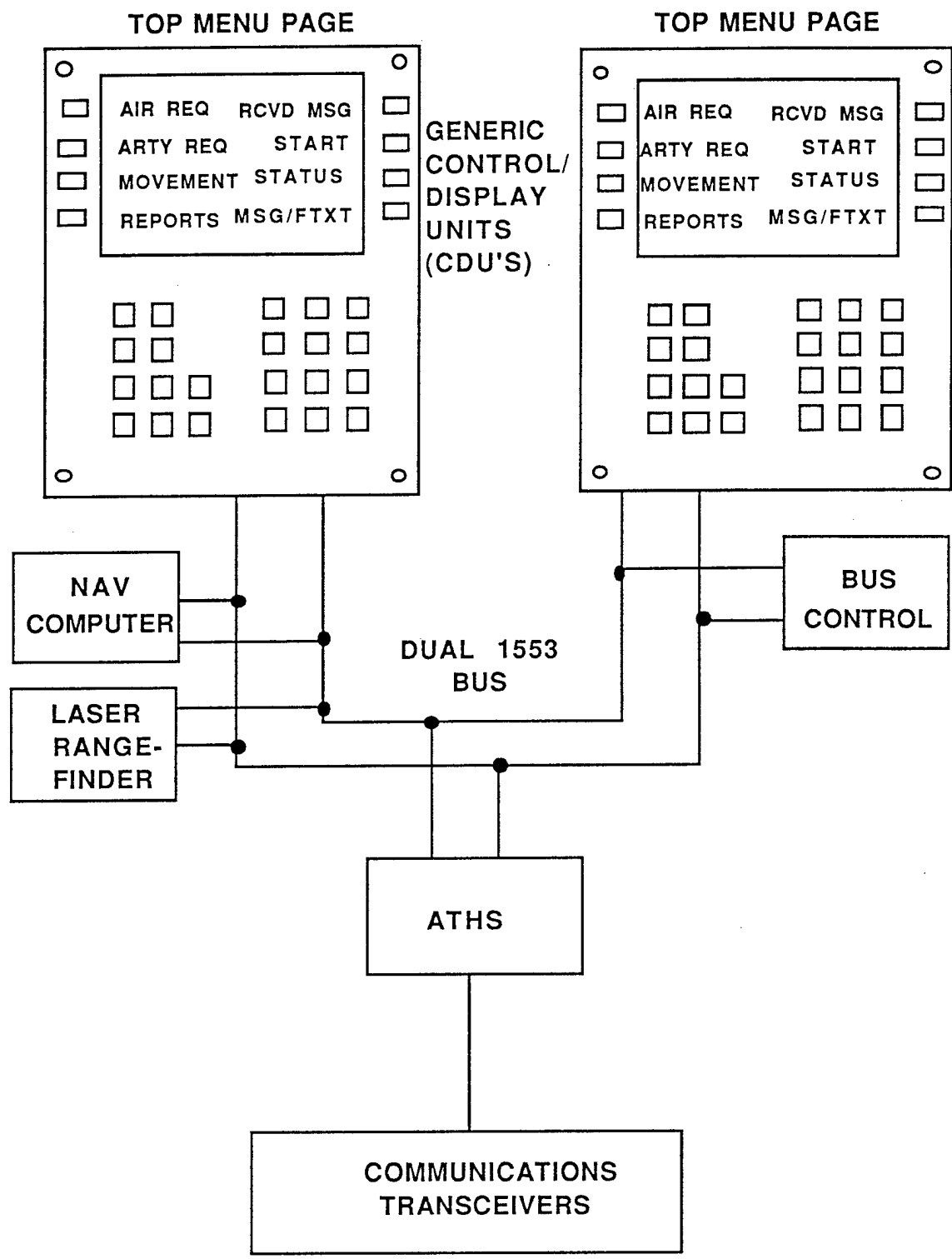
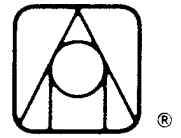


FIGURE 3.1-1: ATHS SYSTEM BLOCK DIAGRAM.



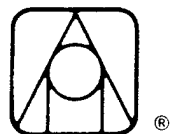
across several Army elements. It follows that our research and results will have generalizability to other future systems although aimed at the AH-64A Apache ATHS/AI.

3.1.1 Description

A typical mission might involve an OH-58D Scout Helicopter and an AH-64 Attack helicopter. The scout detects a target and transmits target data to the attack helicopter. Once the attack element accepts the mission and is in firing position, the scout is notified that all is ready. The scout then moves into position to designate the target and request fire. Following launch, the attack element announces that ordnance is enroute to the target and an automatic countdown to impact is initiated. The attack element is then free either to monitor the current mission or to answer another fire request via the ATHS/AI. The scout sends an "end of mission" message to the attack element when the mission is completed and the target is neutralized.

ATHS/AI provides the capability to simultaneously maintain a number of missions in current mission status. These include up to eight active airborne fire missions, two artillery fire missions, and two preplanned artillery fire missions. The aircrew can also request remote HELLFIRE missions; they can transmit spot, situation, battle damage, and casualty reports.

A typical sequence of actions required to employ ATHS is illustrated in Figures 3.1-2 through 3.1-4. Before performing any of these missions, the operator must initialize the system, which requires several pieces of information. He must enter a one digit code that identifies him as an "Originator". He must also enter a "Team" identifier which specifies the team that the originator belongs to, and the broadcast net on which that team communicates. This code distinguishes the individual aircrew from others on the net. These codes are entered on "Start Page 1" (see Figure 3.1-1). "Start Page 2" provides fields for all parameters of NETS and SUBS (subscribers). These include TACFIRE or AIRNET baud rate, radio, preamble duration, monitor duration, and authentication mode (see Figure 3.1-2). All of this data is required to insure that sending and receiving units can "shake hands" and become synchronized. They can all be entered before the mission begins, that is, before high stress and workload are present to compete for the operator's capacity. An explanation of the subsequent procedures are included in Figures 3.1-2 through 3.1-4 and further described in the following Section (3.1.2).



1. This page is arrived at by pressing the START button on the Top Menu page.

2. Operator selects ORIG and enters a number, selects CURR NET and enters a number, selects BDCST and enters a number. Page 2 is selected by pressing 1/2.

3. NETS and SUBS data pages are accessed by their respective buttons on this page.

4. BULK key allows user to upload automatically all initialization data from another ATHS.

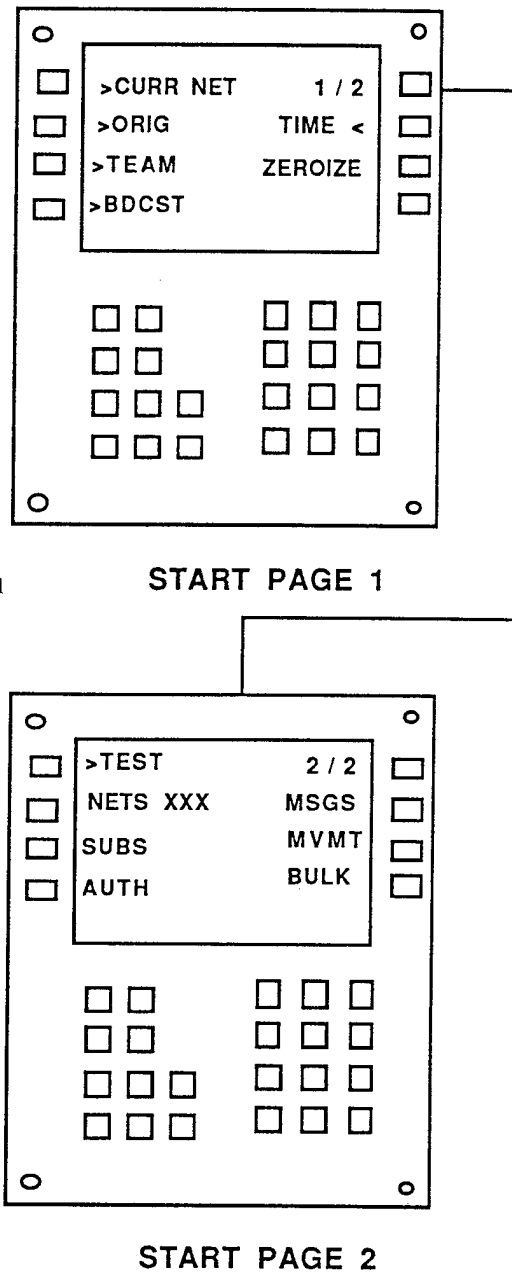
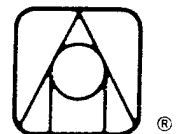


FIGURE 3.1-2: ATHS START PAGES



1. Net Definition pages are accessed by NETS button on Start Page 2. There are eight nets over which the user can communicate.

2. Pressing #PRE and #MON provides access to Preamble and Monitor pages (not shown). These fields must be filled.

The interface shows a central display area with the following text:

AIR NET 2	8 / 8
BLK-SGL	# PRE
BAUD-1200	# MON
RADIO 1	NONE-AUTH

Below the display is a numeric keypad with two columns of buttons:

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Net Definition Page

1. Subscriber Net Assignment pages (3) are arrived by pressing SUBS on Start Page 2. These data are required.

The interface shows a central display area with the following text:

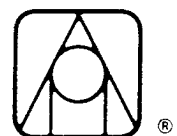
	1 / 3
NET 1	NET 4
NET 2	NET 5
NET 3	NET 6

Below the display is a numeric keypad with two columns of buttons:

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Subscriber Net Assignment Page

FIGURE 3.1-3: NET DATA PAGES



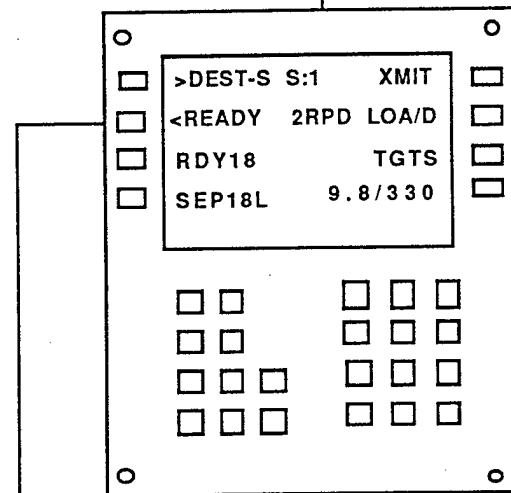
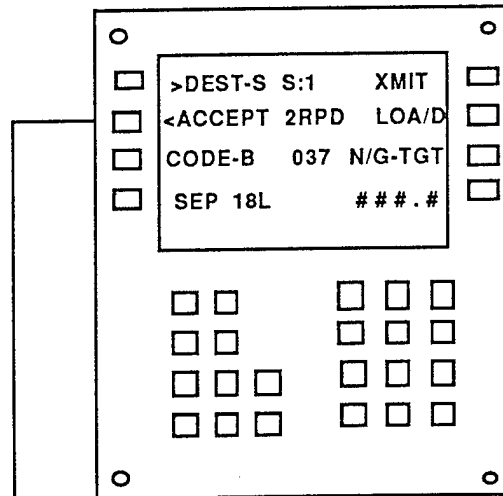
1. Attack aircraft receives fire request with all target data, including range, bearing, type, and ordnance. Attack A/C appears like this.

2. To accept the mission the Attack CP/G presses ACCEPT. The Scout's Mission Summary Page key #3 changes from CODE-B to ACPTD; the Attack A/C Mission Summary Page key #2 changes to READY.

3. When the Attack A/C has maneuvered into position for the shot, the CP/G presses READY.

4. The Scout's Mission Summary Page key #2 changes from READY to FIRE; the Scout presses FIRE and the ordnance is launched from the Attack A/C.

Mission Summary Page

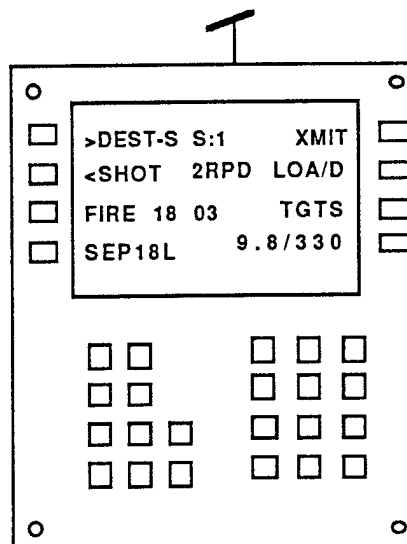


(connect to next page)

FIGURE 3.1-4: PAGE CHANGES AS THEY APPEAR TO THE ATTACK AIRCRAFT DURING "AT MY COMMAND/TARGET" FIRE MISSION.



5. Key #2 on the Attack A/C Mission Summary Page changes automatically to SHOT as the ordnance is launched, initiating an automatic countdown to impact.



6. The Scout then sends an End of Mission message that appears on the Attack A/C Mission Summary Page.

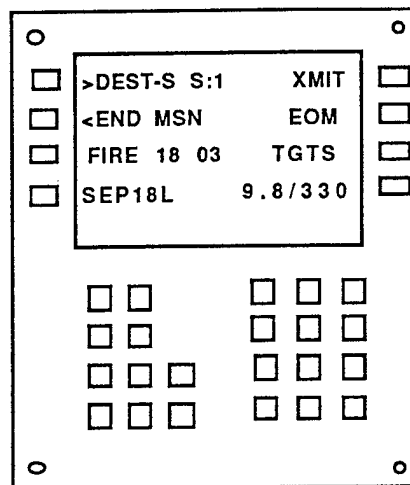
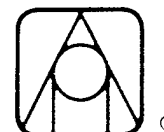


FIGURE 3.1-4: (page 2 of 2) MISSION SUMMARY PAGE AS IT APPEARS TO THE ATTACK AIRCRAFT DURING "AT MY COMMAND" FIRE MISSION.

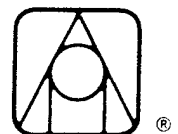


3.1.2 ATHS/AI in Mission Context

Warsaw Pact offensive doctrine calls for massive assaults using armor and mechanized forces. To defeat these numerically superior forces the US Army ground commander will call upon a combined arms force of infantry, armor, artillery, and air power. The Army's air power is organized into Attack Helicopter Companies and Air Cavalry Troops. Their missions are to (1) deny terrain to threat forces, and; (2) perform reconnaissance and communications for the combined arms commander and the artillery Brigade and Division level commanders. Although their missions differ in emphasis, both the Attack Company and the Air Cavalry employ teams composed of attack helicopters and aeroscouts. The aeroscout is tasked to locate targets on the battlefield and provide necessary information to distant artillery emplacements or airborne attack elements to allow accurate weapon/munition delivery. The primary mission of the attack helicopter is *to destroy targets*.

Figures 3.1-2 through 3.1-4 present graphical and narrative descriptions of the sequence of actions required to execute a successful attack mission. The aircraft of interest in the OWL program is the AH-64A and consequently the button pushes presented and the displays generated are applicable to that vehicle working in conjunction with a scout aircraft (e.g., OH-58D). Typically, both the scout and the attack aircraft are concealed from the target and separated from each other by some distance. Of these, the scout determines the location of the targets and can designate them with laser; this guides the munition from the attack aircraft to the target. The actual firing of the missile can be under the control of the scout through various initiation modes (viz., "At My Command/Target", "At My Command/Ordnance" or "Remote HELLFIRE"). It may also be under the command of the attack aircraft (i.e., "When Ready/Target" or "When Ready/Ordnance"). Figures 3.1-2 through 3.1-4 present a typical mission flow for an "At My Command" fire mission.

The CoPilot Gunner (CPG) will normally be the crewmember using the ATHS, although it is available to the pilot. The CPG has a number of duties, including navigation, communication, observer, and sensory system manager. While reducing some parts of the communications responsibilities, utilization of ATHS may be considered either an added or alternative communications task. The added/alternative aspect of ATHS makes it quite amenable to empirical part-task analysis utilizing secondary task methodology (Lysaght et al., 1987). In the context of other tasks, it is unclear that the required sequence of actions

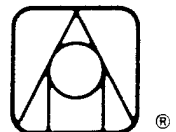


represents low workload, even under the best of circumstances (i.e., good weather, minimal enemy resistance, etc.). However, this is exactly the question which should be asked in OWL analysis.

3.1.3 Previous Efforts Related to OWL

Over the past two years a small but growing body of data relevant to the employment of ATHS/AI in a mission environment have been developed. Even as this report was going into final print, we were still collecting reports directly related to ATHS/AI utilization, as well as detailed task analyses specific to the ATHS/AI/AH-64A combination.

- Of particular value to this program is the development of a detailed task analysis examining cockpit operations of the AH-64 Apache. Anacapa Sciences has developed a particularly fine grained analysis of all control and display activity in the Apache cockpit, under contract to the ARI Field Unit, Ft. Rucker. Additionally, a time-line analysis was included in RFP DAAJ09-87-R-A215 "Integration and Installation of the Airborne Target Handover System/Avionics Integration (ATHS/AI) Onto the Apache AH-64A Aircraft." These will be reviewed, along with alternate task analyses proposed by the successful offeror, to assist in the development of the workload baseline and simulation discussed in Section 3.3.
- Sikorsky Aircraft has also developed a task analysis for ATHS/AI utilization during the HELLFIRE mission on the UH-60A BLACK HAWK. Although this analysis is specific to the BLACK HAWK, it has provided valuable background and experience in communication procedures among firing elements, scout aircraft/forward air controller, and ground-based tactical commanders.
- Another study effort particularly germane to the OWL program consists of a series of flight trials conducted at the Army's Hunter-Liggett location in northern California. Although documentation describing this testing is presently unavailable to us, our understanding is that these flight trials were directed to the examination of operational procedures and relative workload encountered in a light helicopter performing the Scout role. Five aircraft were examined in this role; OH-58C, OH-58C+, OH-58D(AHIP) AH-1S and AH-64. Of particular note, the AH-64 is principally an attack aircraft, but when ATHS/AI equipped, can function in the scout role for sister AH-64s (or any airborne or ground based platform having an offensive capability). We have learned that cockpit activity during these missions was recorded on video tape and presumably analyzed. Taped data are available both for cockpit activity and the outside world as viewed through the gunsight. A test report has been prepared. It is expected that the insights gained during this series of flights will be invaluable for aiding us in the early identification of high workload periods, and particularly the mission related



interrelationships which lead to these intense activity periods. Hopefully, the raw video tapes may be suitable for further analyses for the application and validation of analytic (predictive) techniques such as ProSWAT and ProTLX.

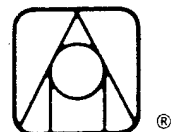
- Additional data which we are in the process of obtaining through Army channels consist of video tape and data analysis conducted on the IBM simulation produced as part of the Advanced Rotorcraft Technology Integration Program (ARTI). The ARTI studies were funded by the Army's Applied Technology Laboratory, and were aimed at investigating the feasibility of single pilot operation of the LHX. Insofar as the LHX was designated for the Scout/Attack role, in addition to a utility variant, an AHS/AI type system would be required. It is our understanding that the IBM simulations included tasking similar to that comprising the AHS/AI mission.

The validation plan discussed in Section 3.3 builds on the available data. The video tapes will be useful for validating the application of analytic techniques. The task analysis data and time-line analysis will be especially useful in developing the operator simulation proposed in Section 3.3 (Objective 2).

3.2 Program Schedule

Figure 3.2-1 illustrates the overall AHS/AI program schedule. The AHS/AI RFP was released on August 1, 1987 and is expected to lead to contract award by the end of December. The program calls for integration and installation on three aircraft, followed by a period of testing. Low rate production commences on successful conclusion of the specified testing.

The schedule for the validation of the OWL Assessment techniques will follow closely the development schedule of the AHS/AH-64 Integration. The OWL validation process is scheduled to require 16 months starting 1 October, 1987. Figure 3.2-2 indicates the schedule for the two major phases with some pertinent milestones from the AHS/AH-64 Integration superimposed. Assuming the AHS/AH-64 contract award occurs on time, all of the requirements of the Validation Plan should be accommodated with one exception. A flyable prototype would be desirable for the assessment involving the empirical technique, however, a prototype may not be available until after our efforts have been completed. A faithful cockpit/mission simulator will be available at that time and can be used for part task empirical analysis.



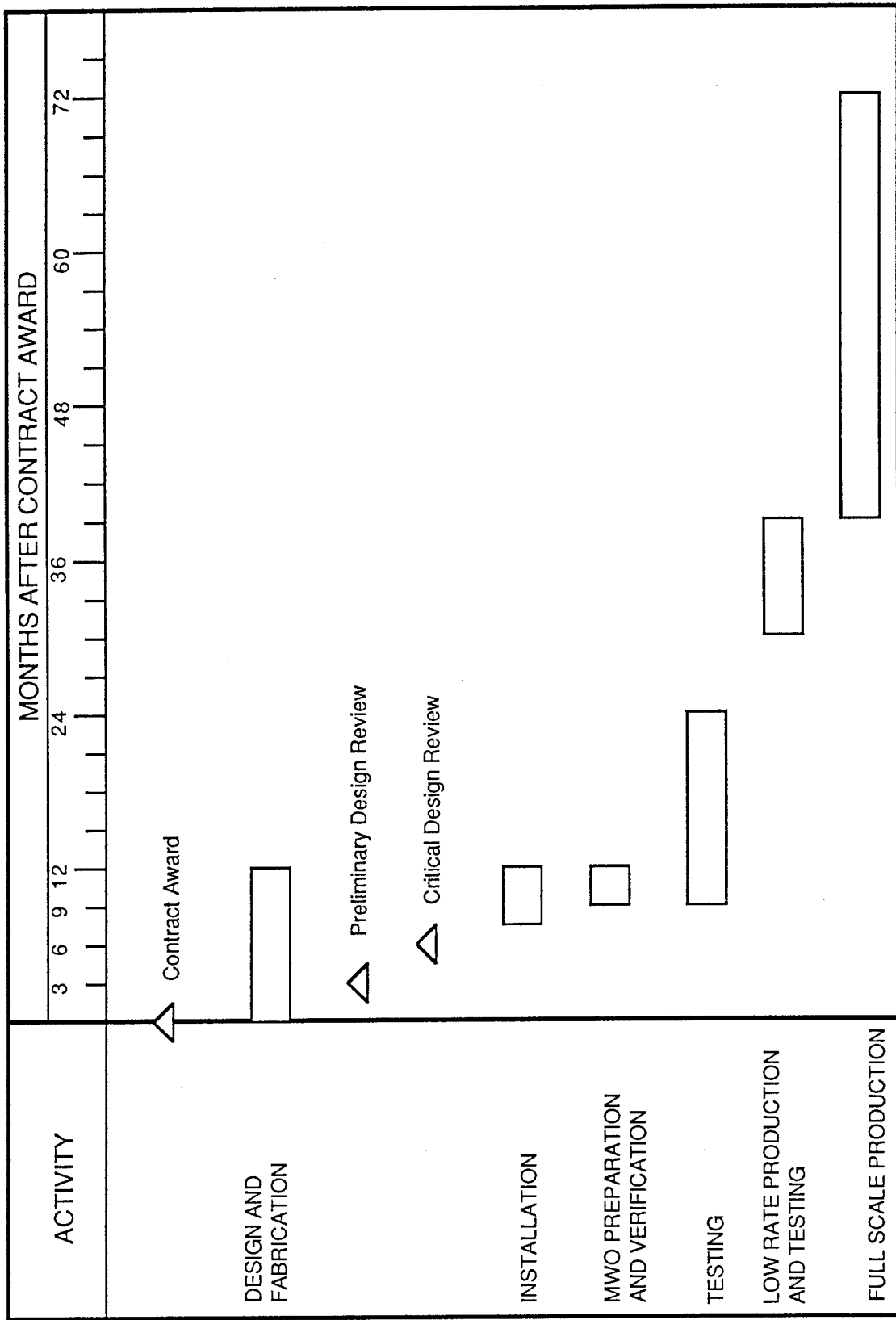
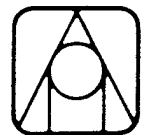


Figure 3.2-1. ATHS Program Schedule



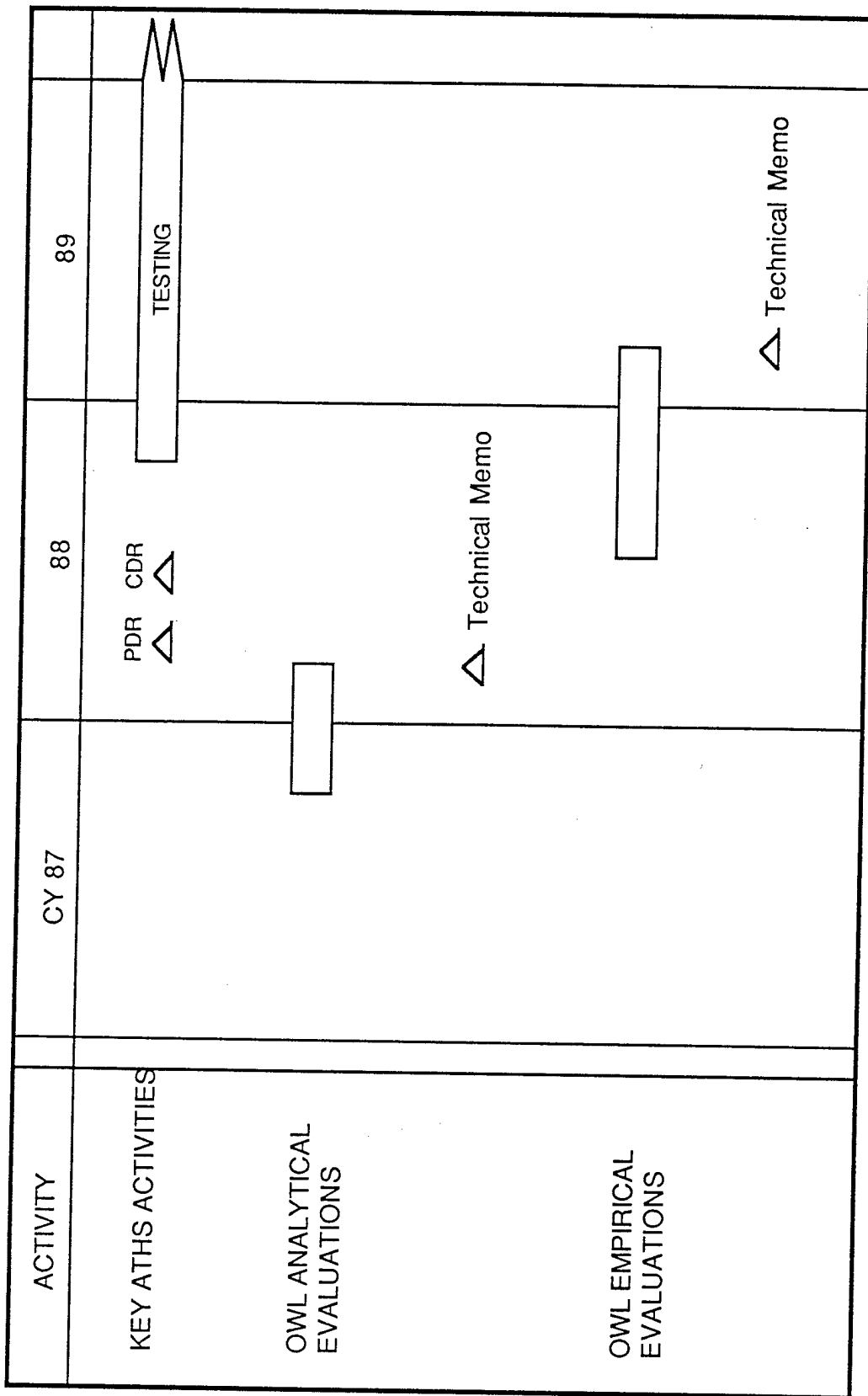
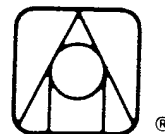


Figure 3.2-2: OWL Validation for ATHS



3.3 Validation Plan

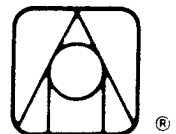
Validity is defined in several ways in the world of testing. The most important of these definitions for OWL is predictive validity: The effectiveness of a test or set of measures in predicting performance in specific situations. The importance of valid OWL measures has been well established in a companion report (Lysaght et al., 1987). Unfortunately, there are no easy or "airtight" tests of validation. We cannot "prove" the validity of a test of a hypothetical construct using another test of a hypothetical construct. Therefore, we must rely on *prima facie* similarities in the results of the various workload measures we have chosen. To the extent that validity builds through progressive iterations of comparison of results and actual workload, validation will be achieved. This is especially true for the analytic (predictive) techniques; these extremely important techniques which have received less validation effort than the empirical techniques.

The validation of the workload measures and techniques applied to the ATHS/AI subsystem will be aimed at three objectives. Consistent with our original thinking on this subject (Lysaght et al., 1987) validation will follow a staged process. Workload evaluation will progress from the analytical to the empirical as the validation phase as seen in the following objectives.

OBJECTIVE 1. Generate a baseline workload assessment for the ATHS/AI through evaluation of system operations independent of a host system.

It is possible conceptually to determine workload imposed by specific ATHS/AI implementations on a variety of host systems while executing clearly defined missions. What is first necessary, of course, is a baseline workload assessment for "pure" ATHS/AI functions. Given a measure of overall ATHS/AI mission workload, it would be logical to assume that the incremental workload above our pre-determined ATHS/AI baseline is due to some combination of the particular ATHS/AI implementation and mission related variables. This is not an argument for strict additivity; it is generally accepted that workload is not an additive quantity. This would appear to be intuitively proper and would provide a practical discriminator among competing designs for a new system.

With this fundamental goal in mind, we will initially develop a workload baseline for the "stand alone" ATHS/AI. This will be accomplished using both prospective methods such as ProSWAT and ProTLX and quite possibly empirical measures collected in a part-

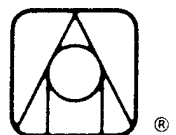


task simulator. We will also attempt to use the video tapes to obtain reflective judgments of workload with SWAT and TLX. Taken together, these measures will permit comparisons of the baseline judgments with judgments on actual performance. It is intended that these estimates shall provide important data for later validation efforts, in addition to providing comparison with other existing baseline analyses

OBJECTIVE 2. Assess operator workload associated with the ATHS/AH-64A integrated weapons system through application and comparison of selected analytic (predictive) techniques.

During this validation phase, the ATHS/AI team will work directly with the cognizant personnel. These will be drawn from the ARI Field Unit at AVSCOM, as well as with other selected subject matter experts familiar with ATHS/AI operations and the anticipated operational environment. A mission scenario will be developed that will provide a standard backdrop against which the results of the analytic and empirical workload measures will be compared. The same mission scenario will be used for analytic and empirical workload assessment techniques. To the extent possible, we will attempt to use the scenario employed in earlier OWL efforts. In this way, continuity can be maintained throughout the validation process. In addition, we will need to define; (1) what the system must accomplish; (2) the time in which it has to be done; and (3) to what level of accuracy it must achieve. These will be used as Measures Of Effectiveness (MOE). Furthermore, time and accuracy data can be developed as primary measures for man-machine system performance.

Pro-SWAT and Pro-TLX, at a minimum, will be applied to the chosen design. Results from these will be contrasted with data collected above in the Objective 1 phase of the plan. More important will be the application during this stage of an analytic modelling technique. At present, the Human Operator Simulator (HOS) is nearing completion of development and in the validation phase for version IV (Harris et al., 1987). This version of HOS introduces a new front end interface and allows the application of HOS on a microcomputer. If HOS IV is judged ready for application to the ATHS/AI validation, then it shall be the model of choice. If application of HOS does not appear feasible then the back-up will be MicroSAINT. This latter model is based upon detailed networking of operator tasks within the context of a mission and defined workstation.



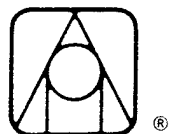
OBJECTIVE 3. Assess operator workload associated with the ATHS/AH-64 integrated weapons system through application of selected empirical techniques as well as validate analytic (predictive) methods.

The current schedules for the OWL program and the ATHS/AH-64 program unfortunately preclude direct workload measurement on the AH-64A flight test vehicle with the ATHS installed. While flight test data would certainly contribute to the ultimate validation, it is not absolutely necessary. Sole reliance upon flight test data would be short-sighted due to the numerous uncontrollable variables. These most common are due to weather interruptions and equipment malfunctions which often reduce the amount of and increase the cost of useful data collected in flight testing. The result may be data insufficient for meaningful analysis.

Empirical techniques can be conducted with a far greater degree of control in a laboratory or simulator environment. However, the laboratory can be a two edged sword. The experimental control is purchased at the cost of the rich mission context afforded by noise, vibration, G-forces, weather variations, and a flight control environment where major errors result not just in a simulator re-start, but in life threatening mishaps. The use of hot mock-ups, part-task simulators, full mission simulators, or integrated bench test facilities all afford the opportunity for meaningful validation of the selected empirical techniques as well as their comparison with the analytical.

Several simulation facilities of the types mentioned above shall be examined by the government/contractor team for our applications. The evaluation criteria will include each facility's capabilities, scheduled availability, and cost of utilizing the facility. The preferences would be in order: an AH-64 simulator, another helicopter simulator, or laboratory work using helicopter tasks. The actual test bed will be determined by availability, cost, and other factors.

Empirical techniques to be employed for Objective 3 will be identified by the OWL matching model, and will likely include the NASA Task Load Index (TLX), the modified Cooper-Harper scale (MCH), and the Subjective Workload Assessment Technique (SWAT). In addition, an embedded task procedure (Lysaght et al., 1987) will be developed in conjunction with the ARI Field Unit at AVSCOM. The empirical data will be used in the validation process for the studies of the analytic measures collected as part of Objectives 1 and 2.

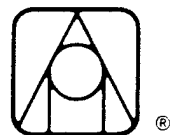


3.4 Summary

The validation approach for ATHIS/AI contains three objectives. The first is to establish a workload baseline, the second is to apply an analytic technique and develop a simulation of the operator, and the third is to perform empirical work to assess workload and validate the analytic techniques used in the first two steps. Table 3.4-1 summarizes the techniques used to accomplish each objective. In line with the iterative nature of OWL evaluation, the phases build on and utilize information collected in previous OWL assessments.

Table 3.4-1: Summary of Measures Used to Accomplish the Three Objectives

Phase Objective	Measures
1	ProSWAT/ProTLX in prospective manner ProSWAT/ProTLX in reflective manner using video tapes from previous work, if possible.
2	ProSWAT/ProTLX Human simulation (HOS or Micro SAINT)
3	Primary measures (time and accuracy) Secondary embedded task SWAT/TLX (use video tape if appropriate)



4. AQUILA REMOTELY PILOTED VEHICLE (RPV) TARGET ACQUISITION / DESIGNATION AND AERIAL RECONNAISSANCE SYSTEM (TADARS)

This section describes the approach to the development and validation of workload assessment procedures and measures for application to the Aquila Remotely Piloted Vehicle (RPV) Target Acquisition/Designation and Aerial Reconnaissance System (TADARS). This section of the report has three subsections. First, the Aquila system is described in Section 4.1. Second, in Section 4.2, the program schedule is delineated, focusing on the opportunities for OWL assessments. Finally, in Section 4.3, the validation plan is presented, consisting of the OWL objectives.

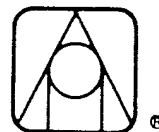
4.1 Aquila System Overview

4.1.1 Aquila Mission Roles

The Army has a long-standing the need for enhanced reconnaissance and target acquisition well beyond the Forward Line of Troops (FLOT). To this end, Aquila is a remotely controlled air vehicle and payload which provides field commanders with a near real time reconnaissance, target acquisition and target engagement information. It consequently fulfills the critical need for near real time target acquisition designation and aerial reconnaissance far into the enemy's area of occupation. It will be capable of providing these functions for both stationary and moving targets.

Aquila will provide enhanced Reconnaissance, Detection and Location (RDL) capability at ranges of greater than five kilometers beyond the FLOT (US Army Field Artillery School, 1986). The designated roles envisioned for the Aquila system are the following (US Army Field Artillery School, 1985):

- **TARGET ACQUISITION:** Detection, identification classification and location of targets in real time. This will permit immediate response by fire support means or engagement at a later time.
- **TARGET DESIGNATION:** Laser designation on both stationary and moving targets for engagement by Precision Guided Munitions (PGM) such as Copperhead, HELLFIRE or terminally guided bombs.



- **ARTILLERY ADJUSTMENT:** Accuracy of target locations provided by Aquila will permit first round fire for effect or one round adjustment on both point and area targets.
- **POST ATTACK ASSESSMENT:** Real time observation of the results of an attack pursued by use of the Aquila system or other systems. Aquila can also video tape results for future analysis.
- **RECONNAISSANCE:** Performs area route and point reconnaissance.
- **BATTLE MANAGEMENT:** Operates within the command, control and communications system at each echelon of employment which permits it to interface with the supported maneuver, fire support and intelligence capabilities. It provides the capabilities described above in real time or near real-time.

4.1.2 Aquila System Description

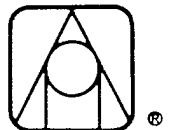
The four major components of Aquila include: (1) the Air Vehicle (AV), with its associated Mission Payload System (MPS); (2) the Launch and Recovery facilities; (3) the Remote Ground Terminal (RGT); and (4) the Ground Control Station (GCS). These components are illustrated in Figure 4.1-1. A brief description of these components is presented below:

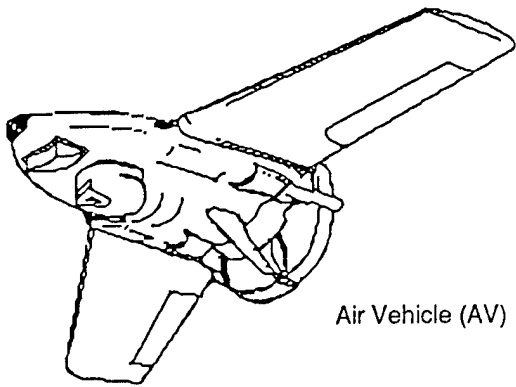
- **AIR VEHICLE (AV)--** the AV is a tail-less, "flying wing" aircraft with a rear mounted engine. The AV is 81 inches long, has a wing span of 153 inches and a height of 31 inches. A pusher propeller is mounted within a circular shroud on the rear of the aircraft. The AV has a gross weight of 265 pounds, a ceiling of 12,000 feet and is capable of air speeds of from 57 to 113 miles per hour. It is capable of flying missions of up to three hours.

The Mission Payload System (MPS) contains a daylight/low light level television camera with selectable fields of view (20, 7.2 and 2.7 degrees), and a 1.06m laser rangefinder/designator integrated with stabilized optics, autotracking and associated electronics. The MPS is powered by the AV power supply. The laser subsystem incorporates a spring loaded eye safe mechanism which requires specific operator action to disengage. Future Aquila improvements will provide a Forward Looking Infrared (FLIR) sensor, providing 24 hour mission capability.

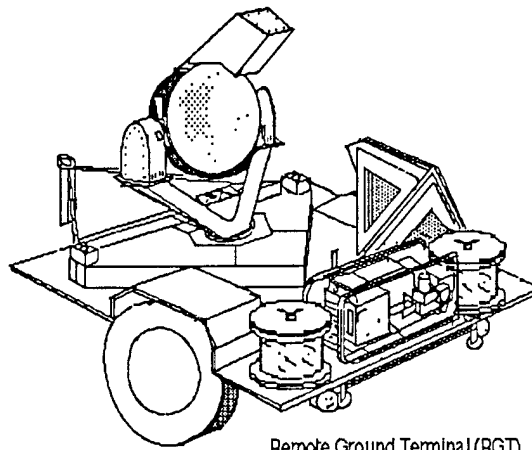
- **LAUNCH AND RECOVERY SYSTEMS --** the Launch System (LS) is a pneumatic/hydraulic catapult mounted on a five ton truck. The system includes subsystems to start the AV, perform preoperational checks of the AV and MPS, and interface with the GCS and RGT.

The Recovery Subsystem (RS) uses a net of vertical dacron straps and decelerator equipment to capture the AV. Navigation into the net is assisted by an integral guidance aid which senses a stroboscopic flasher

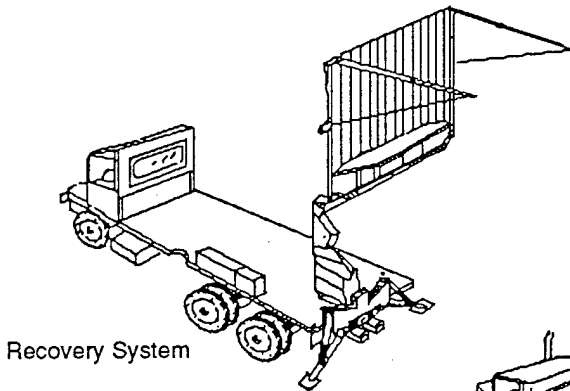




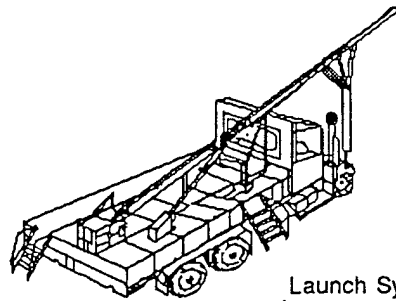
Air Vehicle (AV)



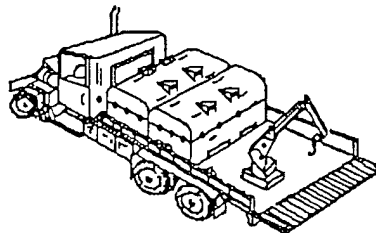
Remote Ground Terminal (RGT)



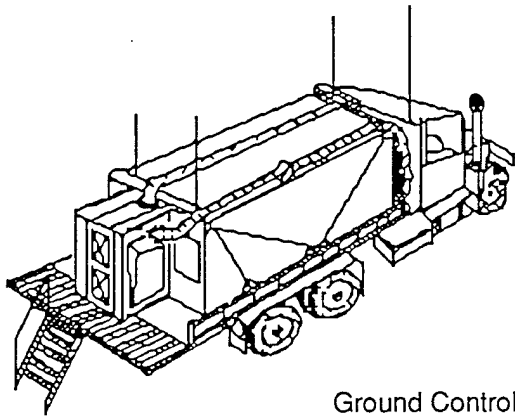
Recovery System



Launch System

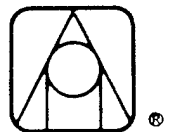


AV Handler



Ground Control Station (GCS)

Figure 4.1-1. Aquila Components



in the AV nose. The AV is captured in a nose down attitude and is recovered from the net using a mobile crane which is integral to the AV storage and transport equipment. An AV handler provides a capability to store and move AVs around the launch site.

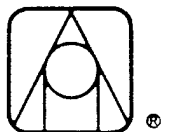
- REMOTE GROUND TERMINAL (RGT) -- this is a trailer mounted microwave transmitter and receiver which provides Modular Integrated Communications and Navigation System (MICNS) data to the AV and GCS. It is remote from the GCS by a pair of fiber optics cables. It tracks the AV, provides AV range, elevation and azimuth data to the GCS, receives telemetry and video data from the AV, and transmits command and control data to the AV. The RGT is powered by a 1.5 KW generator.
- GROUND CONTROL STATION (GCS) -- this contains control consoles for the system and associated communications, command and control equipment and accessories. They are housed in a ballistic and Nuclear, Biological and Chemical (NBC) hardened, environmentally controlled shelter mounted on a five ton truck. The GCS is powered by a 30 KW diesel generator.

4.1.3 Aquila Field Deployment

The Aquila system is presently planned to be fielded in battery-level units. The Aquila battery organization is illustrated in Figure 4.1-2. As shown in Figure 4.1-2, each battery unit consists of a battery headquarters (HQ), two central launch and recovery sections (CLRS), three forward control stations (FCS), and two maintenance sections -- one each for the air vehicle (RPV MAINT) and the ground equipment (GND SPT MAINT).

The CLRS include the five AVs with the associated MPS, the Launch and Recovery facilities (L/R), the RGT, the GCS and associated power generation, maintenance air vehicle handling and transport, and cargo transport capabilities. It launches and recovers AVs for the FCS. It also is capable of independently carrying-out aerial surveillance and target acquisition missions.

The Forward Control Sections (FCS) consist of the RGT and the GCS. The CLRS hand-off the AV to the FCS during initial mission flight segments. The FCS control the actual flight maneuvers and associated data collection during missions, then return the AVs to the control of the CLRS for recovery.



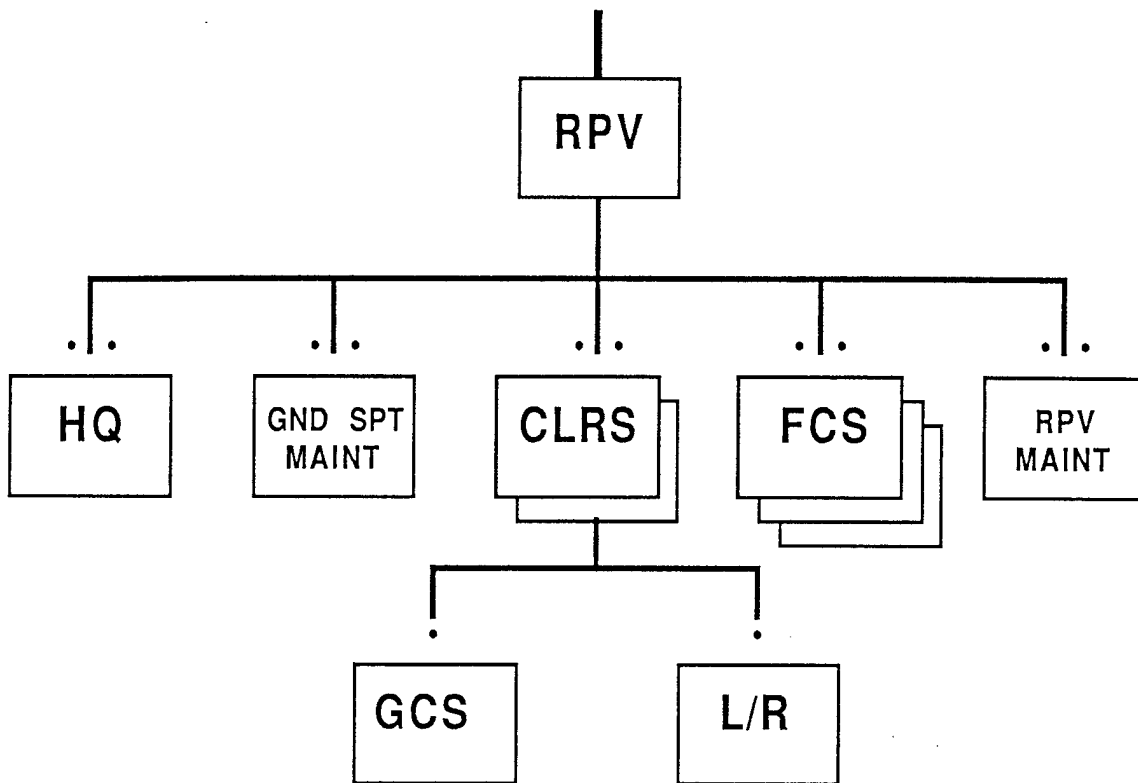


Figure 4.1-2. RPV Battery Organization

Typical battery employment within the division area is illustrated in Figure 4.1-3.

The CLRS are normally employed in the division rear, near the brigade rear boundary. They are typically separated from each other by a distance of four to ten kilometers and located 30 to 40 kilometers from the FLOT. Battery elements are normally co-located with supported unit elements to facilitate security, administration and logistics. The battery headquarters and maintenance elements are normally located with one of the CLRS. The FCS are typically located within the supported brigade area, from 10 to 15 kilometers from the FLOT.

The RPV battery is a Corps asset which is designed to meet division intelligence and target acquisition needs. Support is normally requested through G-3 channels. The corps commander will normally attach the battery to the division. He may, however, retain a portion of the battery for Corps special operations. Missions are requested through intelligence and fire support coordination channels. The G-3 allocates AV sorties in the division operations order. In order to maximize utilization of limited AV assets, the supported brigade S-3 incorporates intelligence, operations and fire support requirements



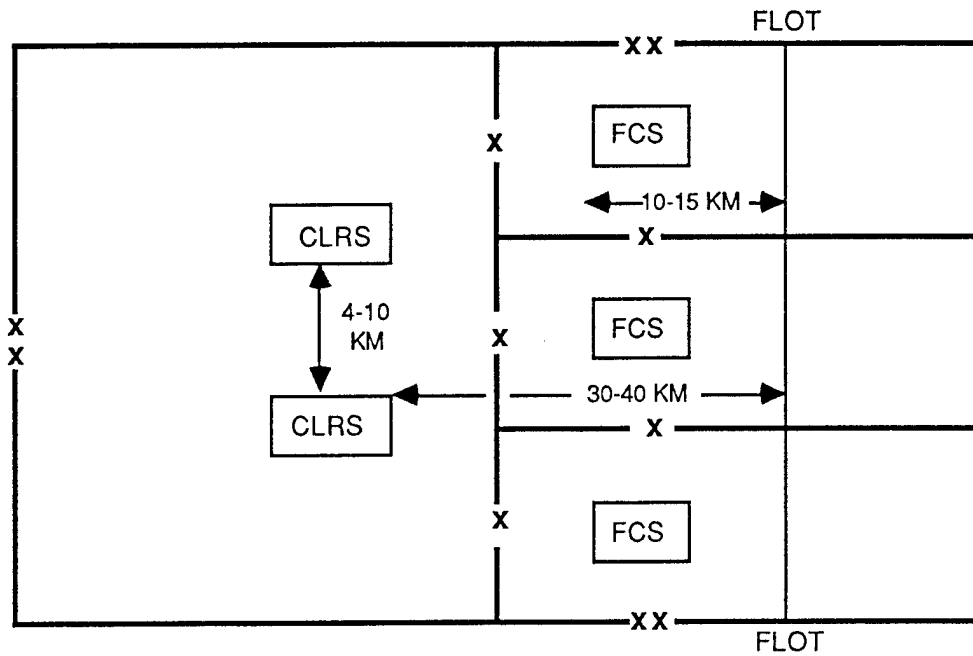


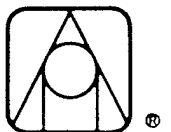
Figure 4.1-3. Typical RPV Battery Employment

into combined missions. As a result, CLRS missions are received from the division Tactical Operations Center. FCS missions for utilization following air vehicle hand-off from the CLRS are received from the brigade Tactical Operations Center.

4.1.4 Conduct of Aquila Missions

The Aquila battery must be capable of successfully performing a variety of tasks in order to perform its mission. These may be viewed as a sequence starting with moving the assets of the battery (i.e., emplacement and march-order), and selection and occupation of CLRS and FCS positions. Following this would be launch operations; hand-off of control of the AV during flight between the GCSs of the CLRS and FCS; and reconnaissance operations. Recovery operations would complete the sequence.

Our efforts for OWL assessment will be directed at the three member crews of the Ground Control Stations (GCSs) within both the CLRS and FCS. It is these crews who perform critical tasks that determine the success of flight missions. The crew members consist of a Mission Commander (MC), a Mission Payload Operator (MPO), and an Air



Vehicle Operator (AVO). To ensure the success of a mission, crew members must perform several main functions.

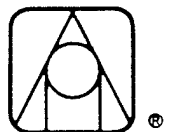
- Upon receiving a mission order, crew members must perform mission planning activities. This includes developing the mission plan (e.g., flight path), and entering the mission plan data, site survey data, and weather data into the GCS main computer. These activities must be successfully completed expeditiously to meet the requirement that the AV be launched within one hour of receiving the mission order.
- During flight operations, crew members must monitor the status and the position of the AV and re-establish the MICNS link, if it is lost.
- When the AV is positioned over each target area, crew members must perform several operations. These include detecting, recognizing and locating targets of military significance. In addition, crew members may be required to designate targets for Precision Guided Munitions (PGM), to call for fires and/or adjust fires as well as to assess damage to targets which have been engaged. RPV functions may be accomplished over several areas during a mission.

Our discussions with the TRADOC and PM communities have indicated that mission planning activities and the subsequent tasks associated with the detection, recognition and location of targets are areas deserving attention with respect to operator workload. The OWL validation plans discussed in Section 4.3 details our approach to addressing these areas. The approach entails the use of families of analytical and empirical OWL techniques in order to identify potential workload problems.

4.1.5 Functional Description of GCS Displays and Controls

OWL assessments will be limited to examining the functions just highlighted within the GCS. The major hardware components of the GCS are shown in Figures 4.1-4 and 4.1-5 and are described below (HQ DA, 1986). It should be noted that the numbers preceding each hardware component description correspond to annotated components in Figures 4.1-4 and 4.1-5.

- (1) COMPUTER/SIGNAL PROCESSING RACK -- this rack contains the system central processing unit and signal distribution and GCS interface components. It is the interface for the AV and MPS operator and Mission Commander consoles. It performs data processing and storage and mission calculation functions.
- (2) MISSION PLANNING FACILITY -- a table top and storage area to support mission planning. An RDS-4502 hard disk drive is mounted under the table top. An AN/UGC 74 Teleprinter is mounted to the right



of the table top. The teleprinter is used to enter mission data and obtain hardcopy system output.

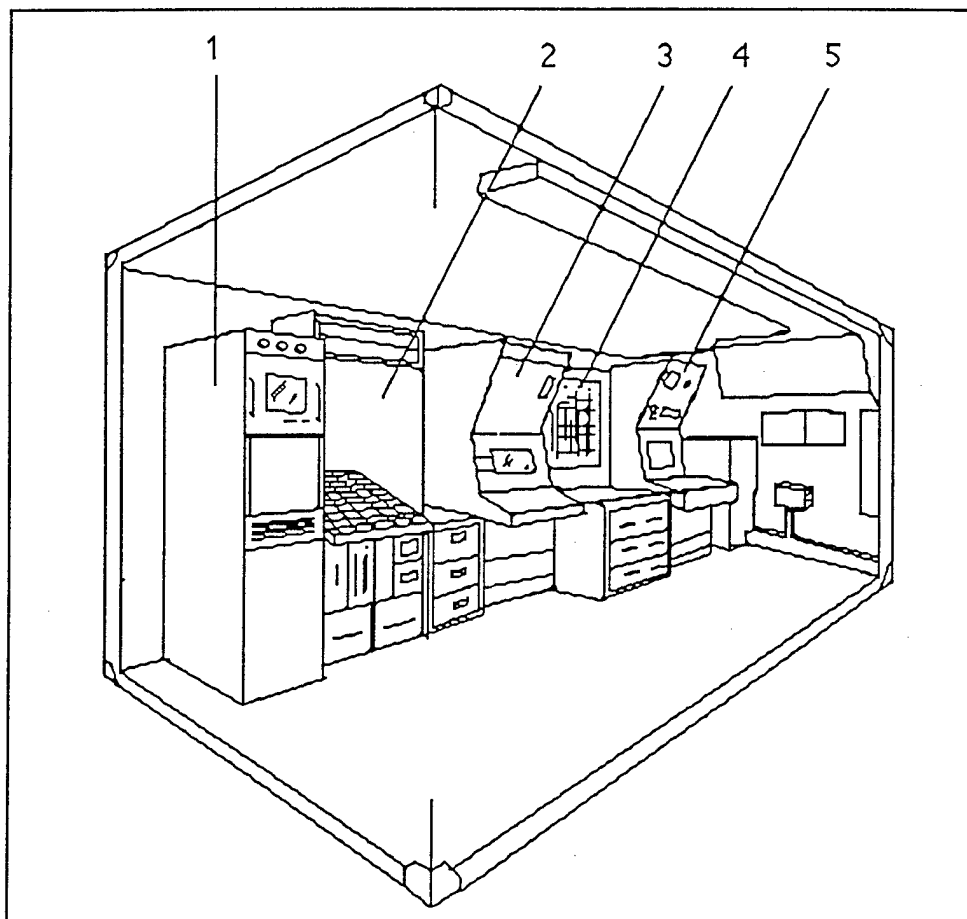
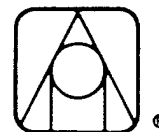


Figure 4.1-4. Internal (Left Side) View of GCS and Components



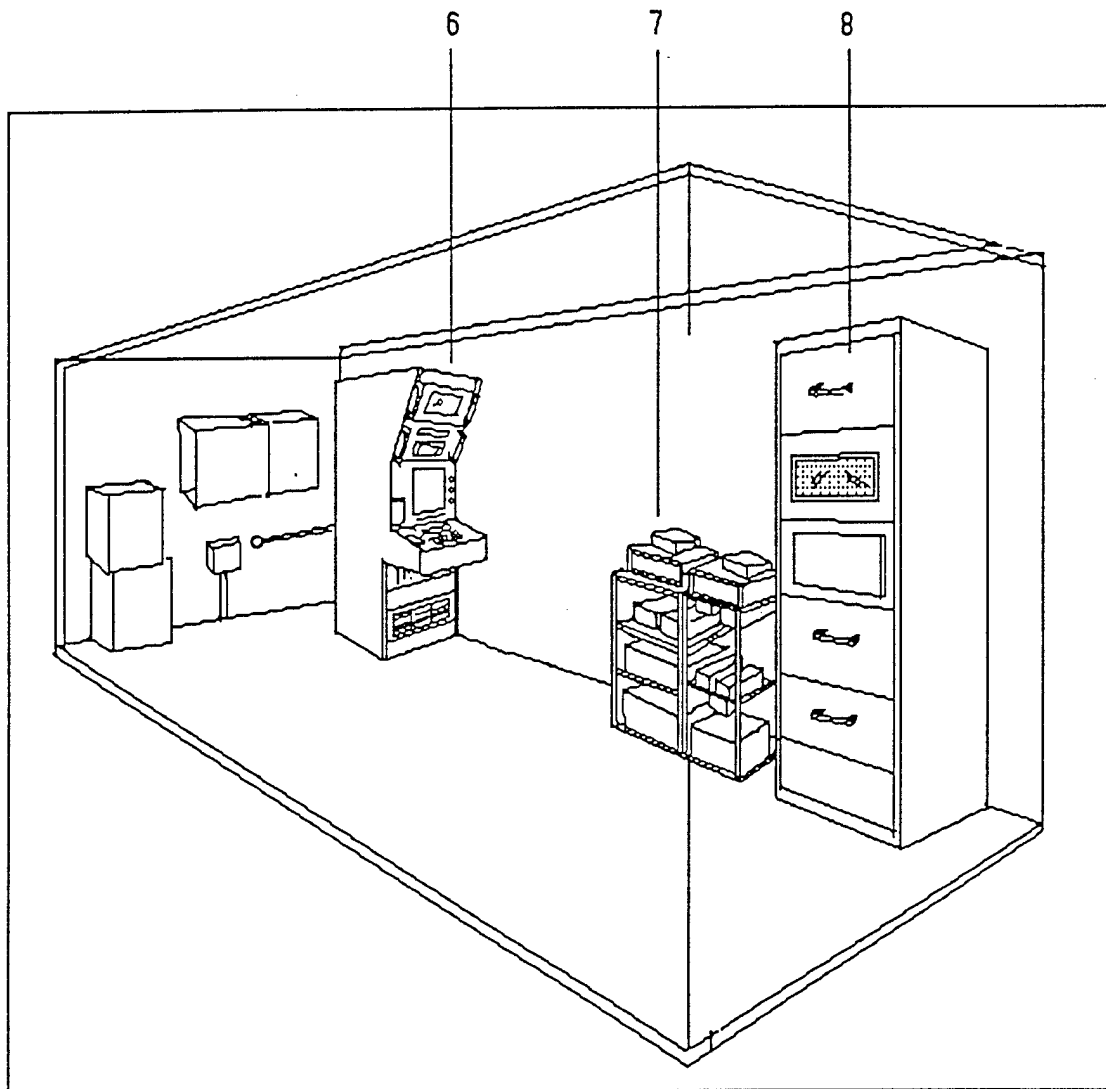
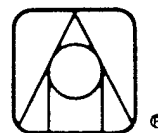


Figure 4.1-5 Internal (Right Side) View of GCS and Components

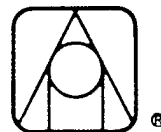


- (3) AIR VEHICLE OPERATOR (AVO) CONSOLE -- this console contains all controls and displays for AV operation from the launch sequence to recovery. A Ground Data Terminal (GDT) control and display panel provides an interface with the RGT. An AV control and display monitors the AV in flight, provides for manual AV control, selects various AV flight modes, and controls parachute deployment (for training) and initiation of launch and recovery operations. A video monitor displays MPS output and AV data.
- (4) NAVIGATION DISPLAY UNIT -- this equipment includes an XT plotter which tracks the AV flight path on a military map.
- (5) MISSION PAYLOAD OPERATOR (MPO) CONSOLE -- this console includes controls for the MPS and displays MPS video output. It is used to select the MPS line of site and field of view, the laser mode of operation (eye safe or eye hazard and laser pulse parameters), and MPS mission parameters (search mode, autotrack, etc.)
- (6) MISSION COMMANDER (MC) CONSOLE -- this console contains MPS laser controls (eye safe or eye hazard), target or landmark cueing controls, a video tape recorder and associated controls and an AN/PSG-5 Digital Message Device (DMD) for entering target data to TACFIRE. It also has a video display of the MPS output. Target data may be superimposed on the display.
- (7) COMMUNICATIONS RACK -- contains standard communications equipment and associated encryption devices for communications with the CLRS, FCS, Battery Operations Center (BOC) and other units.
- (8) DATA LINK RACK -- equipment for interfacing the MICNS between the GCS and RGT.

Not fully annotated on the figures are:

TRAINING INTERFACE UNIT (TIU) -- the TIU may be used to train GCS crew members without the need to employ an AV or aircraft mounted MPS. It may also be used as a reference for maintenance personnel. It simulates MPS and AV output from one of five prepared flight simulations. Equipment faults and data may be entered by the instructor. The TIU is packed in a transit case and is set up in the GCS to the left of the computer/signal processing rack. The unit includes an Imagery Simulation Computer (ISC), associated cables and a Portable Data Entry Terminal (PDET).

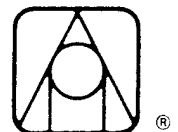
OTHER COMMON EQUIPMENT--Each console contains associated power supplies and communications equipment, including a headset. The headset facilitates communications between section crewpersons and to other locations. The video displays for each console are identical MPS video outputs, with the exception of superimposed data unique to the console function. That data may be suppressed by the console operator. Each video display includes a light pen which may be used to highlight features of interest to other console operators. Light pen output is displayed momentarily on all video displays.



4.1.6 Previous Efforts Related to OWL

The Analytics research team has reviewed all the program documentation which has been made available concerning Aquila MANPRINT and Human Factors Engineering (HFE) efforts. The following summarizes chronologically some of the efforts which have addressed OWL issues:

- A HARDMAN analysis was conducted by the Dynamics Research Corporation for the Army Remotely Piloted Vehicle Program (RPV) (Dynamics Research Corporation, 1983). The HARDMAN analysis generated the RPV Program's projected manpower, personnel, and training demands on the Army's current and/or projected supply of these assets. It did not specifically address any OWL issues or problems. Subsequently, Dynamics Research Corporation performed a re-examination of the support requirements of the RPV system (Dynamics Research Corporation, 1985). In their report, they provided some evidence concerning OWL associated with tasks required to operate the RPV system. Using a modified version of the Subjective Workload Assessment Technique (SWAT), operators were asked to rate the workload associated with specific operational tasks within the context of mission scenarios. This type of prospective workload assessment indicated that the highest workload ratings were associated with the launch task sequence.
- Hughes Aircraft used simulation techniques to study Mission Payload Operator (MPO) performance in simulated copperhead delivery missions (Hershberger, Murray, Fitzhenry, & Farnochi, 1983). Hershberger et al., concluded that the operator could successfully acquire and engage targets, but pointed out a potential decision and task load problem associated with these missions. They suggested a number of control and display modifications in response to this, and other, problems. Crew console modifications made subsequent to that study appear to respond to several of the these suggestions
- Essex Corp, under contract to the ARI Field Unit, Fort Hood, Texas, examined MANPRINT and human factors issues developed during OT-II (Krohn, 1987). The Essex effort centered around 25 critical tasks associated with the operation of Aquila. Reconnaissance, Detection and Location (RDL) issues were not included in the 25. Analytics is currently attempting to obtain the complete task list and will discuss RDL observations with the Essex investigators. Krohn did note problems associated with mission planning activities which could be related to work load. Interruptions, distractions and the time available for mission planning were cited as problems.
- The Aquila System MANPRINT Management Plan (SMMP) (US Army Artillery School, 1987) includes a number of open issues concerning human factors. The copies available to the Analytics team are not completely legible, however, we understand that there are no OWL issues directly addressed. We are in the process of confirming this understanding.

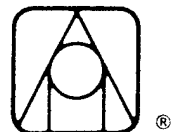


- The Army Research Institute (ARI) among other Army organizations has provided input to the recent Human Factors Engineering Assessment (HFEA) for Aquila. The HFEA was prepared by the Human Engineering Laboratory (HEL) at the Aberdeen Proving Ground (US Army Human Engineering Laboratory, 1987). The HFEA reports human error as suspect for many mission failures during operational tests. Interestingly, the HFEA conclusions infer that the heavy cognitive task load required for GCS crew members is a contributing factor in the number of these errors. The HFEA further reports the need to collect more soldier performance data on critical operations and maintenance tasks. Analytics interviews with Aquila crew members indicated that workload at periods of peak activity, such as during the technical mission planning process, may have contributed to such human errors as cited in the HFEA.
- Efforts are currently underway through the ARI Field Unit, Fort Hood, to examine personnel selection and training requirements for the Aquila system. Such efforts will not directly address OWL issues but may provide valuable information for our OWL assessment effort on Aquila.

4.2 Program Schedule

There are a number of future program options which are currently being considered for Aquila. These range from proceeding with production during fiscal year 1988 to program cancelation. Intermediate options include extending full scale development. The selection of program alternative will undoubtedly be greatly influenced by the performance of the system during the upcoming Force Development Test and Evaluation (FDT&E). The FDT&E is scheduled for October and November of this year. The FDT&E will involve the latest software developed for the GCS facility which may alleviate some the human error problems identified in Operational Test II (OT-II). In this regard, the FDT&E will examine a number of issues that surfaced during OT-II, including crew mission planning and RDL performance.

Figure 4.2-1 illustrates the overall Aquila program schedule. A reliability growth program and development of a FLIR-based MPS are currently in progress. This schedule assumes that FDT&E results and other considerations will support a production decision by the Army Systems Acquisition Review Council (ASARC).



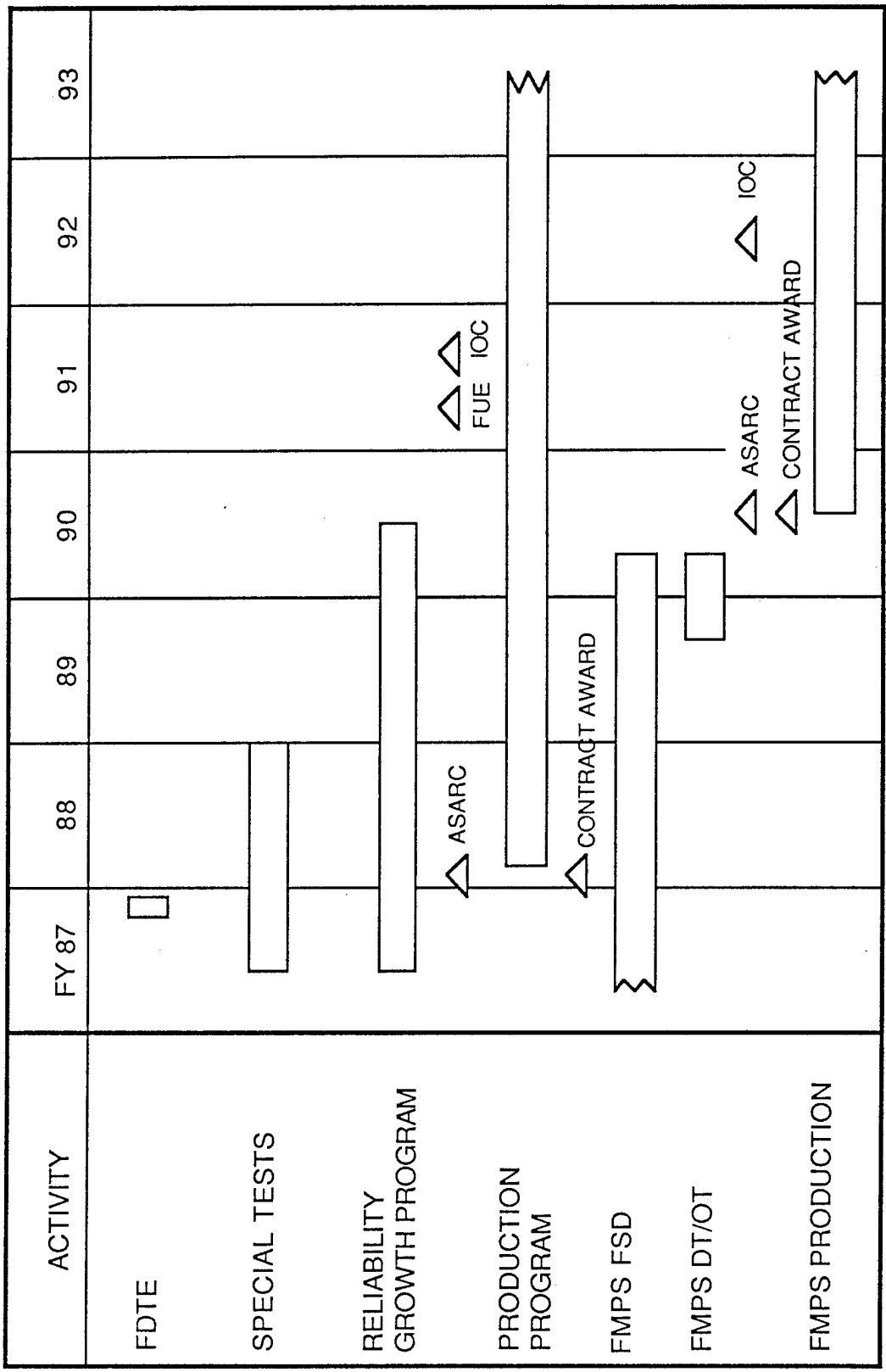


Figure 4.2-1 Aquila Program Schedule

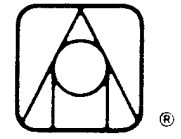


Figure 4.2-2 illustrates the schedule for OWL evaluations of the mission planning and RDL functions, using data collected during previous tests and the FDT&E. Preliminary OWL results from our effort will be made available to the TSM and PM for incorporation into the ASARC package, as appropriate. OWL evaluations for Aquila are scheduled for completion in mid-1988. Technical memoranda describing the analytical and empirical techniques employed and results will be published at that time. The results will be incorporated into an overall OWL assessment technique validation report. That report will be published when all the objectives are completed.

4.3 Validation plan for Aquila

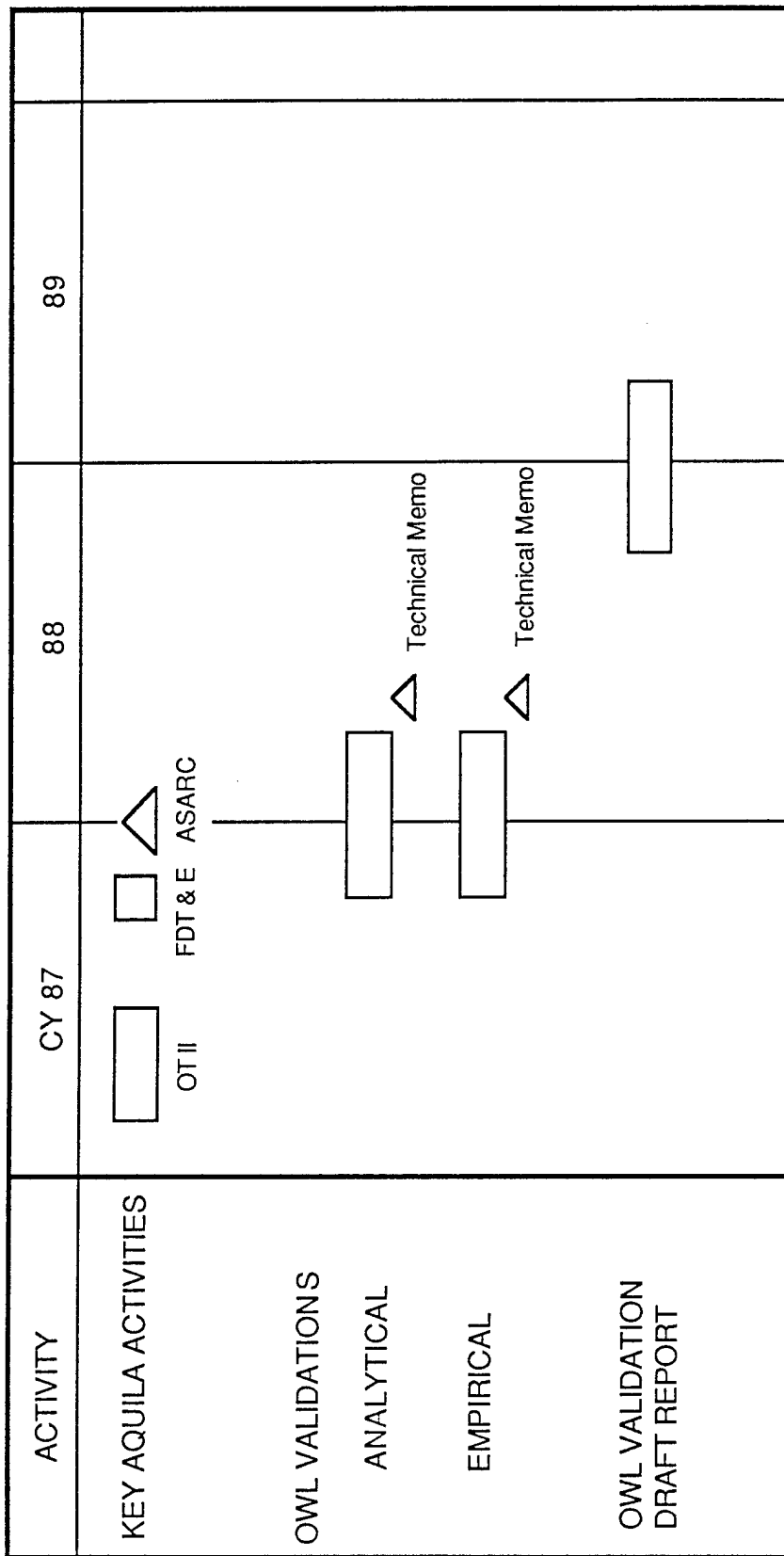
This plan describes our approach to the assessment of workload in the Aquila system with respect to operations of the Ground Control Station (GCS). A more detailed plan fully implementing the ideas espoused in Section 2 will be prepared before the start of data collection. This section consists of a discussion of the objectives of the overall plan and the opportunities that may exist within the Army future test plans for Aquila (e.g., FDT&E) to collect operator performance data.

4.3.1 Objectives of Validation Plan

Given below are four objectives intended to be incorporated into a memorandum of agreement between the various interested parties within the Aquila program and ARI. These objectives represent a merging of the operational needs for OWL information and the scientific capabilities of ARI and its supporting contractors. For each of these objectives, a narrative elaboration of the objective is offered, including description of deliverable OWL assessment products, government inputs required for these products and recommendations for enhancement of mutual efforts.

OBJECTIVE 1. Assess GCS operator workload and the relationship between workload and operator performance by applying a family of empirical OWL measurement techniques to GCS operations (i.e., OT II and/or FDT&E).





△

 ASARC

Figure 4.2-2 Aquila OWL Validations



This will entail OWL assessments of each member of the three man GCS crew by applying a family of empirical techniques (e.g., TLX, SWAT, MCH) during mission planning operations as well as operations associated with the detection, recognition and location of targets. These assessments will be compared to the performance data collected during FDT&E for such operations. Meeting this objective will serve the dual purpose of providing valuable empirical workload assessment, for the new software enhancements being incorporated into the GCS facility **and** offering validation of these workload techniques by comparing their results with actual operator performance data.

This objective will require access to crew members used for FDT&E in order to obtain subjective assessments of OWL. Additionally, these evaluations will require written descriptions of the individual test conditions and any crew member debriefings, along with videotapes of the crew members and their GCS workstations during FDT&E. Also required will be system and individualized timing and error performance data for comparisons with crew member OWL ratings.

The Training Interface Unit (TIU) is being considered in our plans to further explore any chokepoints identified from the effort just proposed. Recall the TIU simulates MPS and AV output from one of five prepared flight simulations. Equipment faults and data may be entered and manipulated by the user. As a result, the TIU provides an opportunity to collect additional operator performance data to verify chokepoints and explore their potential causes

OBJECTIVE 2. Apply a family of analytical (predictive) OWL estimation techniques to GCS operations that were performed during evaluation of the Aquila RPV and compare these analytical assessments with empirical assessments (obtained under Objective 1).

This will require the development of a family of analytical (predictive) individual crew member OWL assessments (via Pro-TLX, Pro-SWAT, etc.). These assessments will be compared and contrasted with empirical assessments obtained under Objective 1. Portions of Objective 4 build upon this validation effort, in that in this objective, we develop a **validated** family of analytical OWL techniques which will be later used to predict the impact of future design changes via P3I.

This objective will required detailed documentation of the Aquila system, scenarios used for the GCS operations (required for Objective 1), and designated "experts" who did



not participate in FDT&E. We are also exploring the possibility to test the "experts" utilized for this objective with the TIU. It may be possible to simulate the conditions employed in FDT&E with the TIU in order to further substantiate "expert" predictive ratings of OWL. This type of testing could further validate the proposed family of predictive OWL techniques.

We are also considering the application of the cognitive analysis approach developed by Zachary and colleagues (Zachary, Zaklad, & Davis, 1987; Zaklad, Deimler, Iavecchia, & Stokes, 1982). Such an approach builds upon a task analysis of mission segments in order to identify the different types of cognitive loading that occur. With respect to mission planning activities, this approach might provide valuable information concerning the apparent problems that have surfaced with this mission segment during previous testing (e.g., OT II).

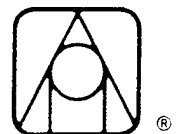
OBJECTIVE 3. Provide results of validated analytical (predictive) and empirical GCS operator workload assessments for potential application to institutional and unit training programs.

This will require letter reports or briefings of emerging results of interest, as well as a copy of the final written report describing in detail the methods and results of the operator workload analyses.

This objective will require access to the personnel data (e.g., ASVAB, SQT) on each operator whose judgments or performance data are used in the analysis. Collection of personnel data will facilitate the examination of individual differences and its potential impact on personnel selection and training for Aquila.

OBJECTIVE 4. Apply analytical (predictive) OWL assessment techniques to potential or planned product improvements for the GCS and estimate their effects on operator workload and performance.

This will entail development of a family of analytical OWL assessments (via Pro-TLX, Pro-SWAT, etc.) of the effects of various potential PIs in the context of scenarios developed in meeting Objective 2. In so doing, the consistency of expert judgments of OWL will be assessed. These judgments will be utilized to estimate the performance impact of reducing OWL by each PI. This objective will require detailed documentation of the proposed PIs as well as access to experienced operators who will serve as "experts".



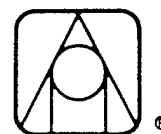
5. LINE OF SIGHT-FORWARD (HEAVY)

This section describes the approach to the development and validation of workload assessment procedures and measures for application to the Line of Sight-Forward (Heavy) (LOS-F(H)) system. The discussion has three parts. First, in Subsection 5.1 the system is described in the contexts of: the overall Forward Area Air Defense System (FAADS) program, the LOS-F(H) system and mission, functional display-control layout, and previous OWL-related assessments. Second, in Subsection 5.2 the program schedule is delineated, focusing on the opportunities for OWL assessments. Finally, the validation plan is presented, consisting of the OWL objectives and a specific empirical design to achieve part of these objectives in Subsection 5.3.

5.1 LOS-F(H) System Overview

A system must be understood in as much detail as is possible in order to develop a plan for the assessment of operator workload. Of course, combat systems are in various stages of the development cycle; the earlier on in this process, the fewer data available. The LOS-F(H) program is an intermediate stage of development, where the selection of the final candidate design prototype will be made in the next several months. Therefore, there is a fair amount of system data available of a **functional**, rather than a specific nature. The purpose of this subsection is to outline and summarize this functional system information as prerequisite to construction of the LOS-F(H) OWL validation plans.

This information is organized as follows. First, the FAADS program is outlined, in terms of both major components and milestones. Second, the available specific information on LOS-F(H) is presented: the mission, system components, and brief descriptions of each of the four candidate prototype systems. Third, a functional description of the display-control configuration is given. This description is based on the US Roland system. Finally, current and previous OWL-related activities are summarized.

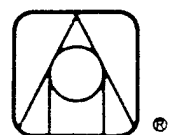


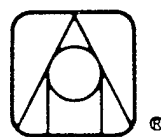
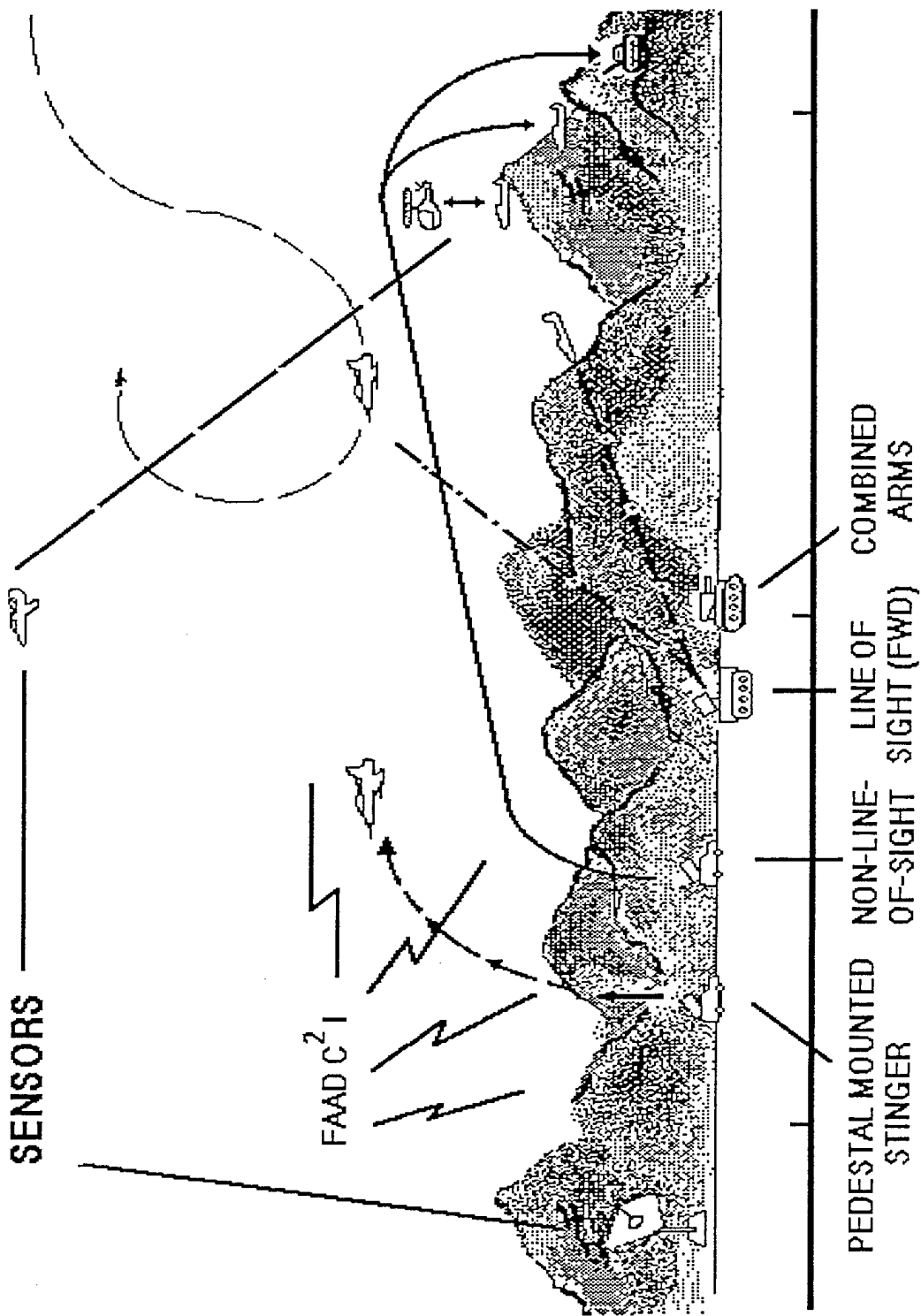
5.1.1 FAADS Program Overview

The LOS-F(H) program is part of the overall development of the FAADS development. In concept, the overall program responds to the FAAD requirements of the Army well in to the next century. Management of the overall FAADS program is the responsibility of the Program Manager for Air Defense/Theater Missile Defense (AD/TMD). Under this management, individual project managers are executing the development and fielding of the various components of FAADS. Those components are being developed either as Non-Development Items (NDI) or under the Army Streamlined Acquisition Process (ASAP).

FAADS is an integrated system. It basically consists of four components: Line-of-Sight Forward (LOS-F), Pedestal Mounted STINGER (PMS), Non Line-of-Sight (NLOS), and Command, Control and Intelligence (C2I). The need to integrate other weapon systems found on the modern battlefield into the FAADS is also recognized. This is the Combined Arms component of FAADS. A pictorial conception of the FAADS and its missions is given in Figure 5.1-1. A brief description of the FAADS components and the associated strategies for developing and fielding those components is presented below.

- LOS-F -- this component of FAADS will provide counter air support to maneuver elements of the division. It must provide a full range of air defense capability in meeting the low altitude air threat which ground maneuver elements face and have mobility and survivability equivalent to the type of force being supported. LOS-F includes LOS-F(H) which is being procured under a competitive NDI approach. Candidate systems are currently undergoing evaluations designed to demonstrate technical and operational capabilities. The LOS-F(H) acquisition strategy recognizes that full "objective system" capability may not be achieved for the system which is initially fielded. The winning candidate will undergo further evaluation during Pre-Production Qualification Tests prior to production buys of the initial "interim system". Full capability will be achieved through a continuous evaluation and pre-planned product improvement approach.
- PMS -- this component will provide defense against low altitude, high performance aircraft operating in the division rear. It will also provide protection against helicopters. It is being procured under a competitive (NDI) approach with the government providing the STINGER launcher as a GFE component.
- NLOS -- the non-line of site component provides a capability to engage slow moving and hovering helicopters which are out of direct line of sight of the air defense system. It will also be capable of engaging ground vehicles. Fibre-Optic Guided Missile (FOG-M) technology





provides this capability. NLOS is being procured under ASAP. Proof of principle and user testing will be followed by competitive procurement orchestrated by an integration contractor.

- FAAD C2I -- the C2I component provides alerting and cueing functions for FAAD assets as well as overall air battle management capability. It is being acquired using multiple NDI acquisitions. A system integration contractor will provide FAAD C2I peculiar software and integrate the Army Command and Control System (ACCS), government furnished communications equipment, vehicles and power sources. The system integrator will also conduct software and system demonstrations.

The general sequence of development for FAADS is shown in Figure 5.1-2. From this figure, it may be seen that all FAAD component development programs are currently in progress. FAAD system capability will be fielded by 1992. Some system components, notably LOS-F(H), may not have achieved full capability at that time under NDI acquisition strategies and will be fielded as interim systems. Full objective capability will be met through application of pre-planned product improvements (P3I).

5.1.2 LOS-F(H) System and Mission

LOS-F(H) will be organic to the FAAD battalion employed in the division area, as pictured in Figure 5.1-1. It is designed to engage helicopters and fixed wing aircraft in the close-in and deep battle areas. In addition, it will have maneuverability and survivability equivalent to that of the supported force. Altogether, it will be capable of providing the following FAAD mission tasks (US Army Missile Command, 1987):

- Engaging helicopters.
- Protecting heavy maneuver forces.
- Attriting threat aircraft.
- Protecting vital assets throughout the heavy division area, in support of the force commanders concept of operation.

The LOS-F(H) program integrates off-the-shelf weapons, fire control, sensors, controls and associated components onto a tracked armored vehicle weapons platform. The system has the following component subsystems:

- Missile(s).
- Missile launcher.
- Gun (objective system).
- Weapons interfaces and interlocks.



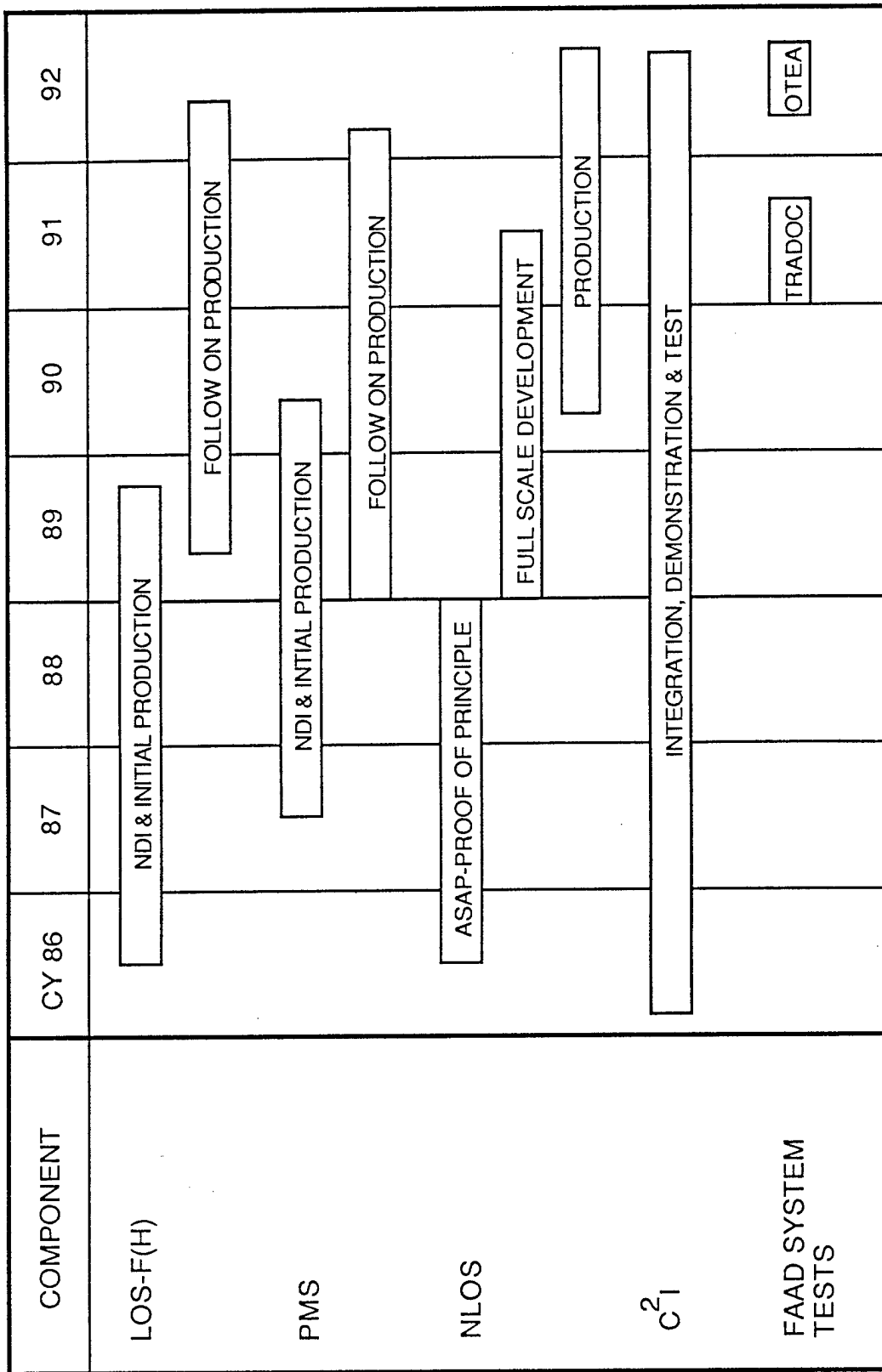


Figure 5.1-2 FAAD Development Program Sequence



- Weapons platforms and elevation and azimuth drives.
- Weapons fire controls.
- Sensors and displays.
- Identification - Friend or Foe (IFF).
- Range finder.
- Optics.
- Controls.
- Communications.
- Tracked Armored Vehicle.
- Support equipment.

LOS-F(H) is designed to engage targets autonomously using integral sensors, optics, range finders and crew observations. Because of this, the interface with the C2I system is an important consideration when evaluating operator workload. This interface must connect LOS-F(H) with a number of sources of distinct types of air defense information, including:

- Track Data -- this data will be received from the FAAD C2I system and may be displayed on a hand-held or remote device. Ultimately, the Enhanced Position Location and Reporting System (EPLRS) will provide this capability. It will also provide automatic slewing to a FAAD C2I directed azimuth through an interface with LOS-F(H) controls.
- Command and Control Information -- LOS-F(H) will use organic radios and the FAAD C2I system processor to maintain contact with platoon/section and battery command posts. Acknowledgment may be made either verbally or via the C2I processor.
- IFF -- a MARK XII IFF system will be integrated into LOS-F(H) to provide IFF functions.

The four systems which are currently undergoing candidate evaluation are off-the-shelf systems. As such, the competing contractors have made a minimum of modifications to respond to the LOS-F(H) request for proposal. Contractors competing in the candidate evaluations have implemented the LOS-F(H) requirement in a variety of forms. Generally, the details of these implementations are considered sensitive and proprietary. Shifrin (1987) described each of the four candidate systems as follows:

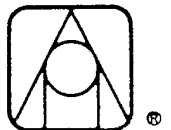
- **Air Defense Antitank System (ADATS)** -- offered by Martin Marietta Orlando Aerospace and Oerlikon Buhrle of Switzerland. A similar system, on an M-113A2 armored personnel carrier, won a competition for the Canadian forces low-level air defense system and is entering production now with first deliveries scheduled for next year.



ADATS, mounted on an M3A1 Bradley vehicle chassis, has a 25mm gun and .50 caliber machine gun and a total of eight laser-beam-riding missiles, with dual-purpose warheads and a range of 8 km (5 mi.). A crew of three operates the system.

- **Liberty Air Defense System** -- offered by LTV Aerospace's Missiles and Electronics Group and Thomson-CSF of France. A derivation of the Shahine system designed for Saudi Arabia from the earlier Crotale air defense system, the future Liberty, with a crew of two, would carry two six-packs of new missiles with folded fins, a derivative of the VT system, and two 25mm guns on a M1A1 tank chassis. The missiles have a range of 10-12 km (6-7 mi.). The version of Liberty to be used for the test and early fielding, an AMX-30 tank with six cannister Shahine missiles, was taken off the French production line. A new radar and fire control system were added.
- **Paladin** -- the air defense system proposed by Western Alliance Air Defense, a joint venture of Hughes Aircraft Co., Messerschmitt-Boelkow-Blohm and Aerospatiale. The earlier fielding version of Paladin incorporates elements of the Roland missile system, including 10 Roland-2 or -3 missiles, mounted on an M-993 multiple-launch rocket system tracked carrier with a 600-shp. motor. A crew of three is required. A second, more advanced version being offered would include a 25mm gun in an anti-aircraft turret mounted on an M1 tank chassis. It would use 10-12 Roland-3 missiles, which have a range of 8 km. (5 mi.), compared with the Roland-2's 6-km (4 mi.) range. The test version uses an XM 975 vehicle.
- **Tracked Rapier** -- offered by United Aerospace Defense Systems, a venture consisting of United Technologies, British Aerospace and FMC - is installed on an FMC/RCM 748 tracked vehicle and carries eight missiles and twin .50 caliber machine guns. The system has been purchased by the U.S. Air Force for defense of U.S. air bases in the United Kingdom. The growth version, called Bradley Rapier, would use the existing production components of Rapier installed on the FMC Bradley armored vehicle, with its 25mm gun. The contractors have installed a Mk. 12 identification, friend or foe capability in the test vehicle. A crew of three is required.

The final relevant LOS-F(H) system information concerns the **specific tasks** that will serve as the focus for workload assessments. For several reasons, OWL evaluations conducted for LOS-F(H) under this project will focus on the **engagement sequence**. First, that task sequence has been studied extensively and is the subject of task analyses which are to be provided by each of the competing contractors. Second, from our discussions with both PM and TRADOC staff, the engagement task is seen to be a potential workload chokepoint. Finally, engagement is directly related to mission goals and is, thus, is functionally very similar across the different systems. Generally, the sequence consists of the following steps:



- Alert Warning
- Target Detection
- Target Transfer (if required)
- IFF Challenge
- Target Tracking
- Counter EW Actions (if required)
- Target Identification
- Engagement Decision
- Fire Decision
- Missile Launch
- Missile Intercept
- Kill Evaluation

There will be slight variations from this sequence based on the candidate system implementations.

5.1.3 Functional Description of the LOS-F(H) Display-Control Layout

An understanding of the configuration of the crewstations and crew member actions during the engagement sequence is essential for preparing the OWL Evaluation and Validation Plan. Candidate selection for the interim LOS-F(H) system will not be made until after October, 1987. It is noteworthy that the Roland Institutional Trainer (RIT), a training and simulation facility in place at Fort Bliss, Texas, has been used to train all crews for the candidate evaluation. This is because US Roland is "functionally" similar in both configuration and operation to each of the candidate systems. We have, therefore, used the US Roland system as a surrogate system for planning purposes. (This use of US Roland is for illustrative purposes and in no way constitutes an endorsement of that system by either ARI or the Analytics team).

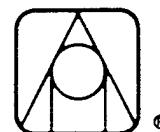
US Roland is currently in the field with the Army National Guard. In particular, one battalion of the New Mexico Army National Guard is equipped and trained with this system. Adequate system documentation is available in the form of Army Field Manuals and Technical Manuals. Additionally, experienced gunners and NCOs are available to provide technical assistance to the evaluation team. We plan to use the RIT to familiarize evaluation team members with typical engagements of the LOS-F(H) experimental application.



The US Roland system is formally known as the **Guided Missile System, Intercept - Aerial: AN/GSG11**. Figure 5.1-3 is a front view of the Roland fire unit in the ground emplacement configuration. The major features of this front exterior view are: (1) Track Radar Antenna, (2) Communications Antennas, (3) Surveillance Radar and IFF Antenna Assembly (facing aft), (4) Launch Beam, (5) Ammunition Magazine, (6) Forward Hatch, (7) Optical Sight, (8) Missile Round. Not shown is the crew entry door located on the rear wall of the of the fire unit. Also not shown is the commanders hatch on top of the fire unit behind the radar antennas. The radars, optical sight and missile launchers are mounted on a rotating turret and cage which contains the gunners crewstation.

The fire unit commander's crewstation is at the rear of the fire unit, facing forward. There is also a forward observers seat and a rear observers seat; however, these are not normal crewstations. Ingress and egress to the fire unit is normally through the rear crew door. Emergency egress may be made through the forward or commander's hatches, unless they are blocked by the position of the turret. Figure 5.1-4 shows the numerically-keyed components of the commander's crew station which are used during the engagement sequence:

- (1) Track Radar Indicator -- provides azimuth, elevation and distance data for targets being tracked. Associated controls include display focus and contrast and control of the type of video input.
- (2) Track Radar Control Panel -- provides frequency and sensitivity controls and antenna scanning and elevation controls for the track radar. The panel also has controls for track radar Built-In Test (BIT) routines and controls for track radar lock-on to a target.
- (3) Surveillance Radar/Pre-Position Indicator (PPI) -- displays surveillance radar output. Data includes target range, IFF engaged, and the position of the missile after launch. Controls include display contrast and brightness, a BIT switch, a selector switch to control whether nearby ground data is displayed.
- (4) Surveillance Radar Control Panel -- provides information on the status of the surveillance radar, selects the radar frequency and sensitivity and selects mode (receive only or transmit/receive). It also provides stage tuning controls and the brightness of ground map video.
- (5) Electronic Counter Measures (ECM) Control Panel -- initiates and selects ECM measures. (Details are classified.)
- (6) Identification Friend or Foe (IFF) Control Panel -- controls for setting and testing IFF equipment. The panel also selects automatic or manual challenge with the IFF system and includes the switch for a manual challenge



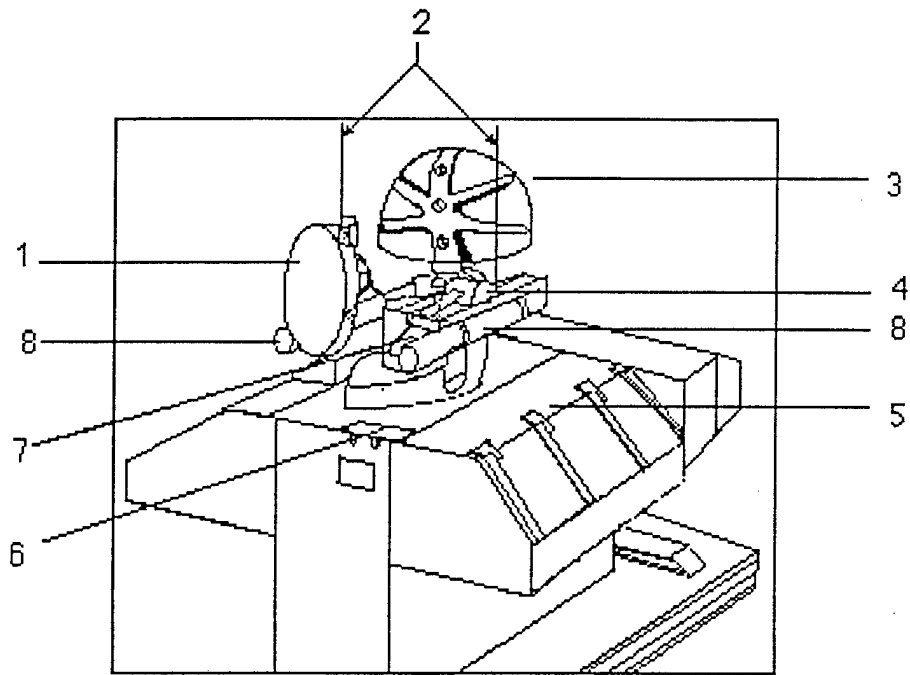


Figure 5.1-3. Roland Exterior View

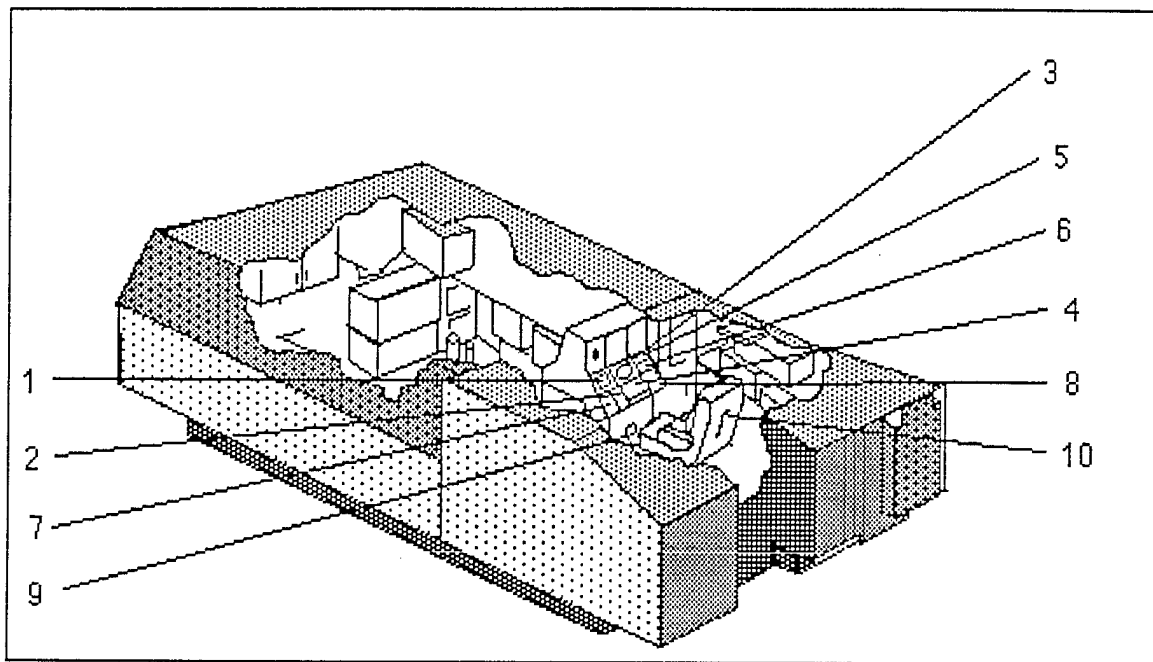


Figure 5.1-4. Commander's Crew Station



- (7) Safety Interlock Control Panel -- controls magazine door operation, and provides indication of an open condition for the rear door and both hatches. Places the turret in the resupply mode (as opposed to normal operational mode), or prevents rotation of the turret ("safe" condition). The panel includes an emergency switch to shut down the primary Power Unit (PPU) and place the fire unit on battery power only.
- (8) Engagement Control Panel -- controls the conduct of an engagement. Includes necessary controls to conduct an engagement with the radar systems, select missiles, eject missile rounds or tubes, and reload the missile rails. Provides status of the conduct of the engagement or reload operation. Includes a switch to terminate the engagement, even if the missile is in flight.
- (9) Commander's Pedal -- fires the missile if the system is under radar control. Allows the gunner to fire the missile if the system is under optical control. Release of the pedal arms the missile in flight. The pedal has a safety interlock which prevents accidental functioning.
- (10) Commander's Seat -- can be adjusted to meet the requirements of various sized individuals. It includes a seat belt and shoulder restraints.

Figure 5.1-5 illustrates the five numerically-keyed principal components of the gunner's crew station which are used during the engagement sequence:

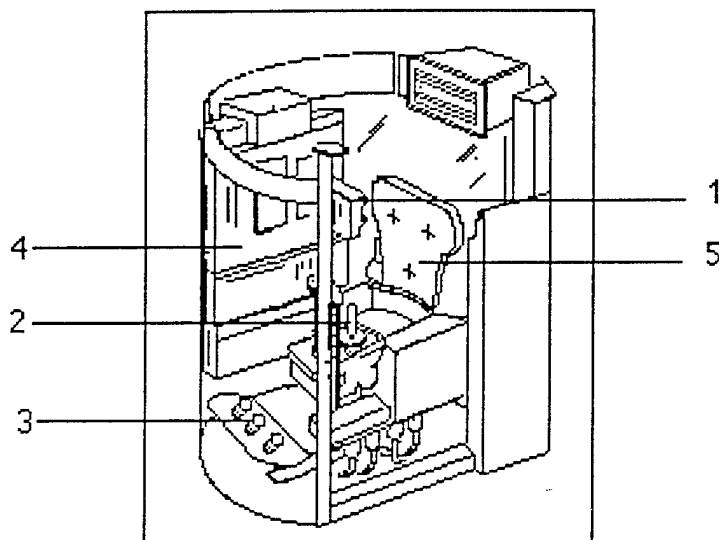
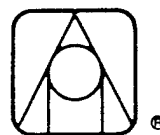


Figure 5.1-5. Gunner's Crew Station

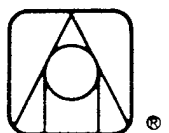


- (1) Optical Sight and Reticle -- provides the optical sighting capability for the system. Focus, eyepiece distance and the headrest are adjustable to meet the gunner's needs. The reticle permits the gunner to align the system on the target and provides information concerning whether the system is under radar or optical control. The reticle also signals if the commander has pressed the commander's pedal to fire the missile or to enable the gunner to fire.
- (2) Gunner's Control Panel -- a control stick is used to control the azimuth and elevation of the launch beams, thus, the line of site of the optics. It has a height adjustment to accommodate different sized crew members. Other controls include a missile command transmitter channel select switch, reticle lighting controls, and filters for the optics. A terminate switch terminates the engagement in the same manner as the commander's terminate switch. An optics/radar switch slaves the optical site field-of-view to the track radar, or permits tracking independent of the radar. A proximity fuze disable switch disables the missile proximity fuze.
- (3) Gunner's Pedals -- the gunner uses three foot pedals. One is used to obtain sight control from the commander and change the field of view of the optical site (178 or 89 mils). Pressing the pedal the first time takes control from the track radar; subsequent operation of the pedal switches field of view between wide and narrow. The second pedal controls elevation and azimuth tracking speed. Maximum speed is obtained with the pedal fully depressed. That pedal is operated with the gunner's right foot. The third pedal fires the missile if authorized by the fire unit commander.
- (4) Built-In-Test (BITE) Equipment Panel -- provides status of the fire unit (e.g., sight/radar control, BITE results) and controls for using BITE. An emergency off switch disconnects all power, except battery power, and shuts the PPU down in the same manner as does the commander's emergency off switch.
- (5) Gunner's Seat. The height of the gunner's seat and the distance to the controls are adjustable to accommodate various sized crew members. The seat includes a seat belt and shoulder restraints.

Communications within the fire unit are maintained with an intercom system using standard Combat Vehicle Crew (CVC) helmets and control boxes. Crewmen may also use external radios. The external radios are controlled by the fire unit commander.

5.1.4 Current and Previous Efforts Related to OWL

Workload-related issues are recognized during the NDI development of LOS-F(H). Cited in the LOS-F(H) System MANPRINT Management Plan (SMMP), they are being examined during the candidate evaluations and are to be addressed during follow-on



testing. The following summarizes some of the current/near-term actions which address OWL issues:

ARI, Fort Bliss, Texas is supporting the LOS-F(H) MANPRINT program with a variety of activities and services. ARI personnel and supporting contractors have made substantial contributions to the overall LOS-F(H) MANPRINT program and are supporting the candidate evaluations. These contributions include the preparation of studies and documentation which serve as a foundation for OWL activities. Among them are conduct of MANPRINT-related Front End Analyses (FEA) to include job and task analyses, support to the Source Selection Evaluation Board (SSEB) for LOS-F(H) and test support during the candidate evaluations.

No specific OWL-related analysis is required from the candidate system contractors. However, each of the four contractors has agreed to furnish a Critical Task List, which details the different operator tasks required by that particular system to accomplish its missions, focusing on the engagement sequence. These task lists will form the foundation for a task analysis.

ARI supporting contractors are to evaluate and integrate these candidate system Critical Task Lists into a Mission Task List. The Mission Task List will be a generic task analysis -- applicable to each of the candidate systems -- describing the sequences of operator tasks and subtasks required to perform a given mission, focusing on engagement. The Mission Task List will also provide a basis for both training and personnel requirements analyses.

ARI and its supporting contractors will be collecting system operability and Human Factors Engineering data throughout the candidate evaluations. This data will include:

- engagement sequence timing data on tasks and subtasks,
- error data on frequency, type, and recoverability, and
- videotapes of each crewmember of each system performing mission tasks, including instrument operation.

These performance data offers excellent opportunities for the validation of both analytic and empirical workload measures.

The Aldridge-McCracken methodology will be applied to crewmembers of each candidate system to produce task difficulty ratings for each task and subtask of the Mission



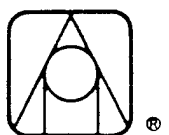
Task List. The difficulty ratings and the engagement sequence timing data (obtained from the candidate evaluations) will be combined with the Mission Task List. The resulting workload analysis will be reported to the SSEB.

5.2. Program Schedule

The LOS-F(H) acquisition sequence is pictured in Figure 5.2-1. The candidate evaluation is currently in progress, consisting of prototype testing of the four candidate systems described above in Subsection 5.1.3. Selection of the winning system will be made in the latter part of CY 1987, based on the candidate evaluation process. Following this decision, the winning system will undergo the Phase I FDT &E in the Spring of 1988. In the Phase I FDT & E, two reconditioned models of the winning candidate will undergo technical testing, which will eventually support a production decision for the interim LOS-F(H) system. The first delivery of 10 firing units will be followed by the Phase II FDT & E and Initial Operational Test and Evaluation in fiscal year 1990. Subsequent production options provide for delivery of an additional 238 firing units through Fiscal year 1994.

The early operational tests will permit collecting data and examining issues not addressed in the candidate evaluations. The LOS-F(H) Program Manager has indicated that these first follow-on tests will be with equipment which still has many of the "warts", which will be removed from the eventual fielded system. For instance, several of the candidate evaluation systems do not have adequate gun capability to fill missile system dead space. Another example may be the requirement that any single crew member be capable of successfully engaging a target. Most of the "warts" will be resolved through design changes implemented for the interim production system and evaluated in the second phase of the test program. Modifications to bring LOS-F(H) to full objective system capability will be addressed in P3I implemented later in the life cycle.

This acquisition approach provides an excellent opportunity to develop, apply, and validate OWL evaluation products which can favorably influence the final LOS-F(H) design. Figure 5.2-2 shows the proposed schedule for OWL evaluation activities and for delivery of OWL products in support of LOS-F(H). For our program, the two critical activities are the current candidate evaluation, and the Phase I FDTE. Predictions made from a selected battery of analytic OWL techniques will be developed using the generic



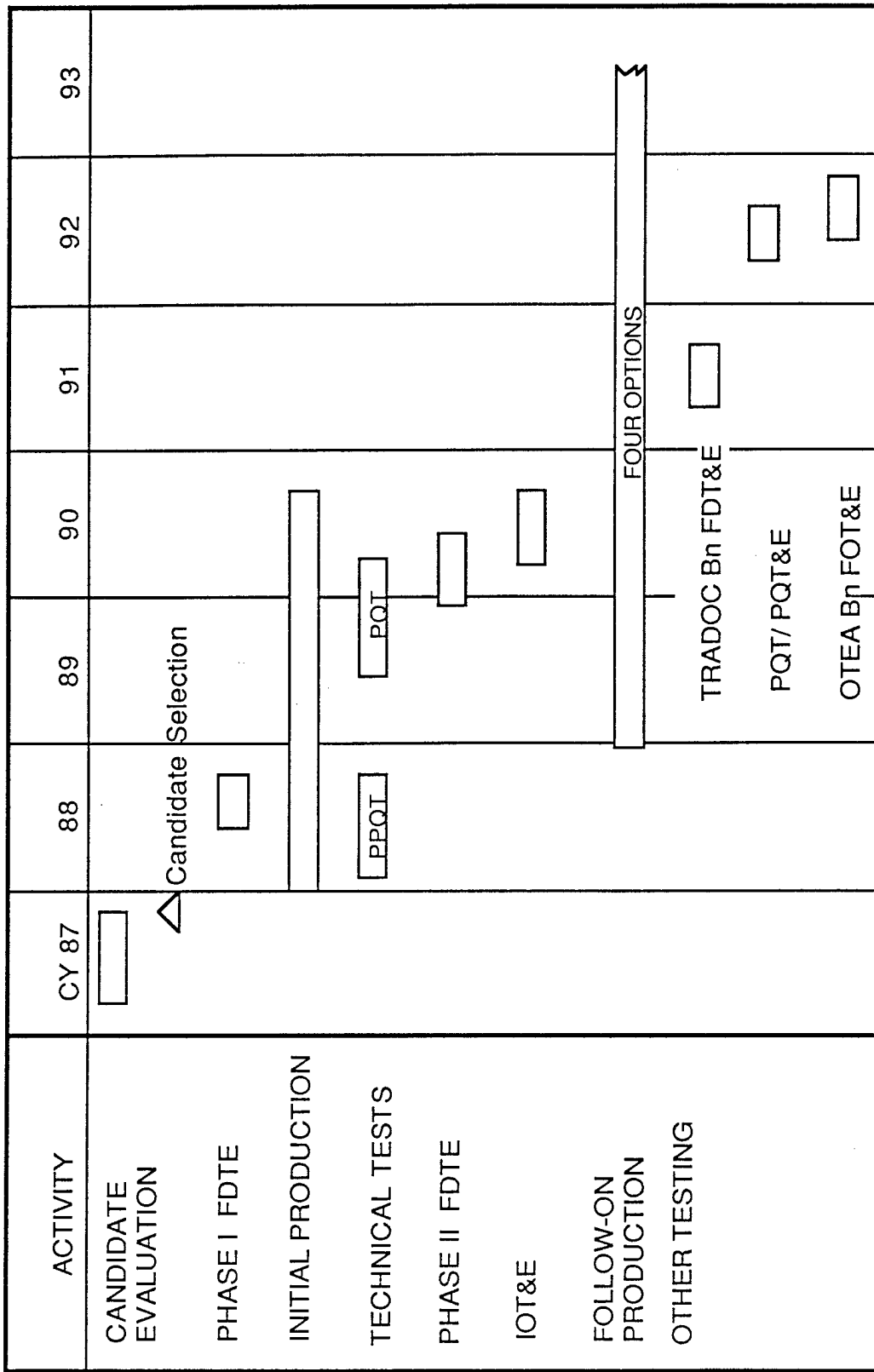
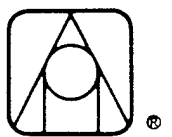


Figure 5.2-1 LOS-F(H) NDI ACQUISITION



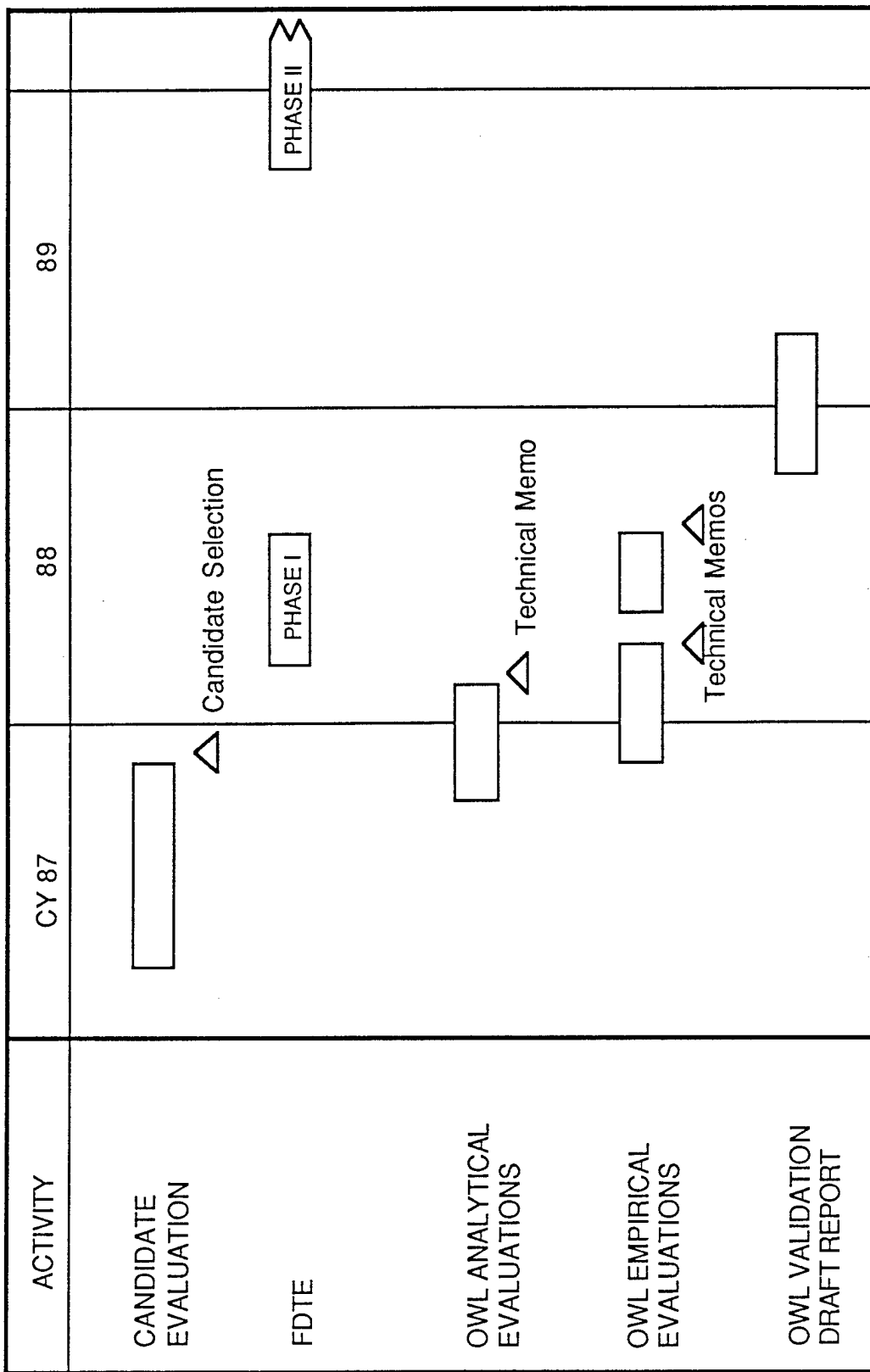
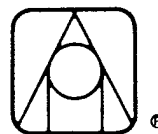


Figure 5.2-2 LOS-F(H) OWL Evaluations



Mission Test List (itself based on the candidate contractors Critical Task Lists). These predictions will target the engagement sequence and the activities of each of the crew roles. The predictions will be validated from the candidate evaluation using empirical workload techniques. Such data on the winning system will become available in early Calendar year 1988.

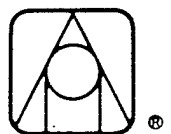
Later in 1988, the Phase I FDT &E will provide an opportunity for further predictive and empirical workload evaluations. We will have the "golden opportunity" to make substantive input to the **design** of the FDT & E tests in areas related to the evaluation of operator workload. By that time, we will have analyzed the candidate evaluation data and have validated a battery of analytic techniques. This knowledge will be reflected in the FDTE design. Our objectives and plans for utilizing these OWL assessment opportunities are discussed below in Subsection 5.3.

5.3. Validation plan for LOS-F(H)

Now that the LOS-F(H) system is understood in terms of its mission, its component subsystems, and its acquisition strategy, a sound plan for OWL assessment and validation can be constructed. This plan delineates our approach to the identification and diagnosis of potential workload problems in the eventual fielded LOS-F(H). We have already provided an approach to the **analysis** of workload data in Section 2. In this subsection, an illustrative plan is provided for the **collection** of such data. The plan uses all available knowledge and resources as efficiently as possible -- to get the "biggest bang for the buck". To do this, the plan must have clear and relevant goals, must take advantage of all possible assessment opportunities, and must have a design that is sound from both a scientific and an operational point of view. This subsection details the proposed validation plan for the LOS-F(H) system. It consists of a discussion of the objectives of the plan and an exemplary experimental design to begin the empirical evaluation of those objectives.

5.3.1 Objectives of Validation Plan

Given below are four objectives, intended to be incorporated into a memorandum of agreement between the various interested parties within the LOS-F(H) program and the Army Research Institute. These objectives represent a merging of the operational needs for



OWL information and the scientific capabilities of ARI and its supporting contractors. For each of these objectives, a narrative elaboration of the objective is offered, including description of deliverable OWL assessment products, government inputs required for these products, and recommendations for enhancement of mutual efforts.

OBJECTIVE 1. Assess crewmember OWL and the relationship between OWL and operator performance by applying a family of empirical (subjective) OWL techniques to operators of the winning candidate system.

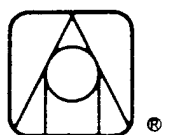
This will result in commander and gunner assessments of OWL by applying a family of empirical techniques (e.g., TLX, SWAT, etc.) to engagement sequences. These assessments will be compared to the performance data collected during the candidate evaluation, when this data becomes available. Meeting this objective will serve the dual purpose of providing valuable empirical workload assessments for the winning candidate system and offering a comparison of their results with the accuracy and timing data.

This objective will require access to crewmembers of the winning system for empirical subjective assessments of OWL. Additionally, these assessments will require written descriptions of the individual test conditions and any crewmember debriefs, along with videotapes of the crewmembers and their workstations over engagement sequences during candidate evaluations. Also required will be system and individualized timing and error performance data for comparisons with crewmember OWL ratings.

It is recommended that empirical OWL assessments be obtained from every crewmember in each of the three crewstations. By doing this, we could begin evaluation of individual differences, especially when applied to the stringent system conditions and greatest expected levels of OWL. In particular, we recommend comparing individual performance data with individual subjective assessments, as well as comparing means. Collection of personnel data will facilitate the examination of individual differences (e.g., ASVAB, SQT).

OBJECTIVE 2. Apply a family of OWL analytic (predictive) techniques to the winning candidate system and validate them against empirical assessments.

This will result in development of a family of predictive crewmember assessments (via Pro-TLX, Pro-SWAT, etc.). These assessments will be compared and contrasted with independently derived task-analytic McCracken-Aldrich Model estimates and with empirical



assessments obtained under Objective 1. Objective 3 and portions of Objective 4 build upon this validation effort, in that in this objective, we develop a **validated** family of predictive techniques which will be later used to predict the impact of future P3I design changes.

This objective will require detailed documentation of the winning system, scenarios of the engagement sequences (required for Objective 1), and designated "experts" who did not participate in the candidate system evaluations. Comparable McCracken-Aldrich results as well as performance data (also required for Objective 1) are also needed.

It is recommended that predictive evaluations be obtained from every crewmember in all three crewstations. Such data could provide for initial contrasting of the analytic and McCracken-Aldrich Model assessments in the context of individual differences. Validation of these two types of techniques would be achieved by comparing these estimates with the performance data collected in reaching Objective 1.

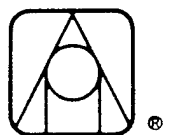
OBJECTIVE 3. Apply family of analytic OWL techniques to potential Product Improvements (PI) for the winning candidate and estimate the effects of PIs on OWL and operator performance.

Using the validated analytic instruments developed under Objective 2, this objective will result in predictive estimates of the effects of various potential PIs in the context of scenarios developed in meeting Objective 2. In doing so, the consistency of expert judgments of OWL will be assessed. These judgments will be utilized to estimate the performance impact of reducing OWL by each PI.

This predictive assessment objective requires detailed documentation of the PIs as well as access to winning system crewmembers and/or other system experts for collection of judgments. McCracken-Aldrich estimates could also be considered for the purpose of this evaluation, if modeler expertise and other resources were made available.

OBJECTIVE 4. Fully incorporate use of OWL measurement techniques during the LOS-F(H) FDTE-Phase I.

The Analytics team will participate in the design of FDT & E Phase I, using a family of analytic OWL methods (e.g, Comparability, Pro-TLX, Pro-SWAT) to provide guidance in the selection of test conditions. In addition, we will provide crewmember



assessments by applying a family of empirical OWL techniques (e.g, TLX, SWAT, MCH) as part of the evaluation of engagement sequences in the FDT & E Phase I. These OWL predictive and empirical results will be compared with performance data obtained during the testing. The comparisons will provide guidance and refinement of the predictive assessments of product improvements developed under Objective 3.

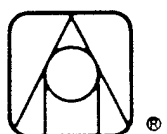
This objective will require access to winning system crewmembers and other system experts before the Phase I FDT & E tests in order to exercise the predictive instruments. Comparability analysis will be conducted if practicable under time and other constraints. During the conduct of the FDTE, access to the crews of the LOS-F(H) will be required for the empirical assessment of OWL. These assessments will require participation in crewmember debriefs and access to videotapes of the crewmembers and their workstations over the engagement sequences. For the comparison of the OWL predictive and empirical results with performance, FDT & E performance data will be required.

LOS-F(H) issues of one- and two-man-operability and interoperability are important OWL considerations. To study these factors, the FDT & E design must allow for fire unit testing with varying size of crew and with crewmembers changing roles. It is also recommended to study the effects of training and the effects of individual differences, both independently and in conjunction with the other main variables.

5.3.2 Experimental Design for LOS-F(H)

The following is an outline of an illustrative experimental design for the accomplishment of one portion of Objective 1 (i.e., empirical (subjective) assessment of OWL and comparison with performance data for the candidate evaluation tasks). It is based on our current knowledge of the details of the crews, missions, and structure of the candidate evaluations. There will, of course, be inaccuracies and changes with respect to the actual conduct of the tests. Thus this design is a **proposed** design, subject to the constraints of the operational environment.

Overall, the design proposes to administer three subjective OWL measures to each of three crews of three persons each, performing ten distinct tasks or missions. The main variables studied will be crew, subject within crew, mission, and workload measure



(corrected for artifacts of mission order or measure order). These empirical results will be analyzed and compared with the candidate evaluation performance data as indicated in Objective 1.

The design outlined below is not necessarily limited to Objective 1. In fact, it could also be used with the predictive versions of the OWL measures. For example, both the experience/training levels of the subjects and the level of detail of the system/mission descriptions could be varied, and thereby studied as main effects or interactions.

5.3.2.1 Subjects

It is anticipated that three crews, each composed of three crew members (i.e., driver, gunner, and commander) will participate in the study. *At this point, we believe that the winning system will provide two fire units, each with two trained four-man teams.* Each of these four teams is to consist of a commander (E-5) and three gunner-drivers (E-2 to E-4). We have built our design on the conservative assumption that only three out of four crews will be available for testing.

5.3.2.2 Missions

It is anticipated that each crew will have participated in ten selected missions or tasks. Mission descriptions will be examined and each mission will be rated for its probable difficulty (or overall crew workload). One of the ten missions will be selected as the "benchmark" mission. The remaining nine missions will be ranked from "1" (lightest) to "9" (heaviest), ordered by expected crew workload. The nine missions will be further subdivided into three groups of missions, each group containing one mission each from the bottom, middle, and upper thirds of the rankings. The three groups of missions and the sequence in which the crews will be exposed to them are shown in Table 5.3-1.

As can be seen in Table 5.3-1, the average ranked difficulty is held constant across the three crews so that mission order can be studied. The benchmark mission will be administered to all crews both before and after the nine ranked missions. It will be used to determine the presence of possible dynamic effects (e.g., learning, fatigue) across the missions.

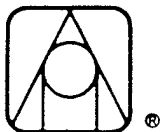


Table 5.3-1 Illustrative Experimental Layout

	ORDER OF MISSIONS									
	1	2	3	4	5	6	7	8	9	Sum
Crew #1	1	5	9	6	7	2	8	3	4	45
Crew #2	6	7	2	8	3	4	1	5	9	45
Crew #3	8	3	4	1	5	9	6	7	2	45
Sum of Ranks	15	15	15	15	15	15	15	15	15	135

5.3.2.3 Workload Measures

For each mission, each subject will be administered three subjective OWL measures including SWAT and TLX. The sequence in which subjects respond to the three OWL measures will be systematically varied from mission to mission. Both the sequence in which the OWL measures are administered and the administration time will be recorded. Administration time will be used to subsequently establish the relative cost-effectiveness of the three OWL measures used. Sequence of administration will be used as a main effect to determine the impact on rated workload brought about by having previously estimated workload by an alternative OWL measure.

5.3.2.4 Data and Analysis

Both objective measures and evaluator ratings of crew performance will be collected and compared with measures of subjective workload derived from the three OWL measures. The purpose of this comparison is to investigate the utility of aggregated individual crew members' OWL estimates for predicting team performance.

Analysis of Variance (ANOVA) and other techniques will be used to identify the main variables' (and their interactions') contributions to; (a) rated workload; and (b) performance data:



- Crews
- Subjects within Crews
- Missions
- Order of Missions
- Workload Measures
- Sequence of OWL measures

The relationship between means and standard deviations of workload estimates will be computed. The relationship among Subject effects and MOS scores will be obtained to determine if those effects are predictable by the MOS score. This should provide insight into the impact of individual differences on subjective ratings of workload and the extent to which they may be predictable. The relationships among performances measures and workload measures will be obtained to determine the extent of overlap in the workload measures and actual or rated crew performance. Analyses will follow the formulation presented earlier in Section 2.



6. CONCLUSIONS

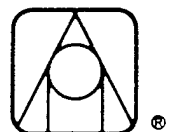
6.1 OWL Analysis

The measures scheduled to be applied to each system have been enumerated in the individual sections. Figure 6.1-1 provides a summary of the OWL methods used across systems. Examples of all method categories are used except analytic/math models and empirical/physiological. Unfortunately, the systems do not offer an opportunity for these two categories of measures. The reader may wish to refer to Lysaght et al. (1987) for a comprehensive discussion of the various categories of measures. Primary empirical measures are used on all systems as are the cost-effective and quantified opinion measures. These later measures consist of SWAT, ProSWAT, TLX and ProTLX. SWAT and TLX have received considerable attention in the context of validation and serve in part as a benchmark for the analytic techniques.

Some of the measures are "new" in the form proposed and represent expansion of the set of validated measures available. For example, ProTLX has not been employed in any previous studies. Further, other techniques such as Comparison and Simulation Models have not been well-validated against empirical data. We propose using cognitive task analysis to expand into an increasingly important area of OWL concern (Zachary et al., 1987).

6.2 Management Plan

We have discussed in Sections 3, 4 and 5, our validation plans for the three systems. Our management approach to implement these plans involves three semi-autonomous research teams coordinated administratively by John Bulger. The team approach provides the flexibility which is necessary due to the geographical spread of the three sites as well as the diverse nature of the systems themselves. Figure 6.2-1 shows the



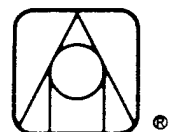
organization as well as the team membership. Alvah Bittner will head the Aquila Team in addition to his technical oversight roles as Program Manager. Each of the team leaders has had prior experience with systems similar to his responsibility.

<u>METHOD</u>	<u>CATEGORY</u>	<u>ATHS</u>	<u>AQUILA</u>	<u>LOS-F(H)</u>
Analytic	Comparison			●
	Math Models			
	Expert Opinion	●	●	●
	Task Analytic Methods	○	●	○
	Simulation Models	●		
Empirical	Primary Task	●	●	●
	Operator Opinion	●	●	●
	Secondary Task	●		
	Physiological			

● OWL Analysis To Be Done in Validation

○ OWL Analysis Previously Done

Figure 6.1-1. Taxonomy Annotated with Workload Assesment Techniques Used on Each of the Systems



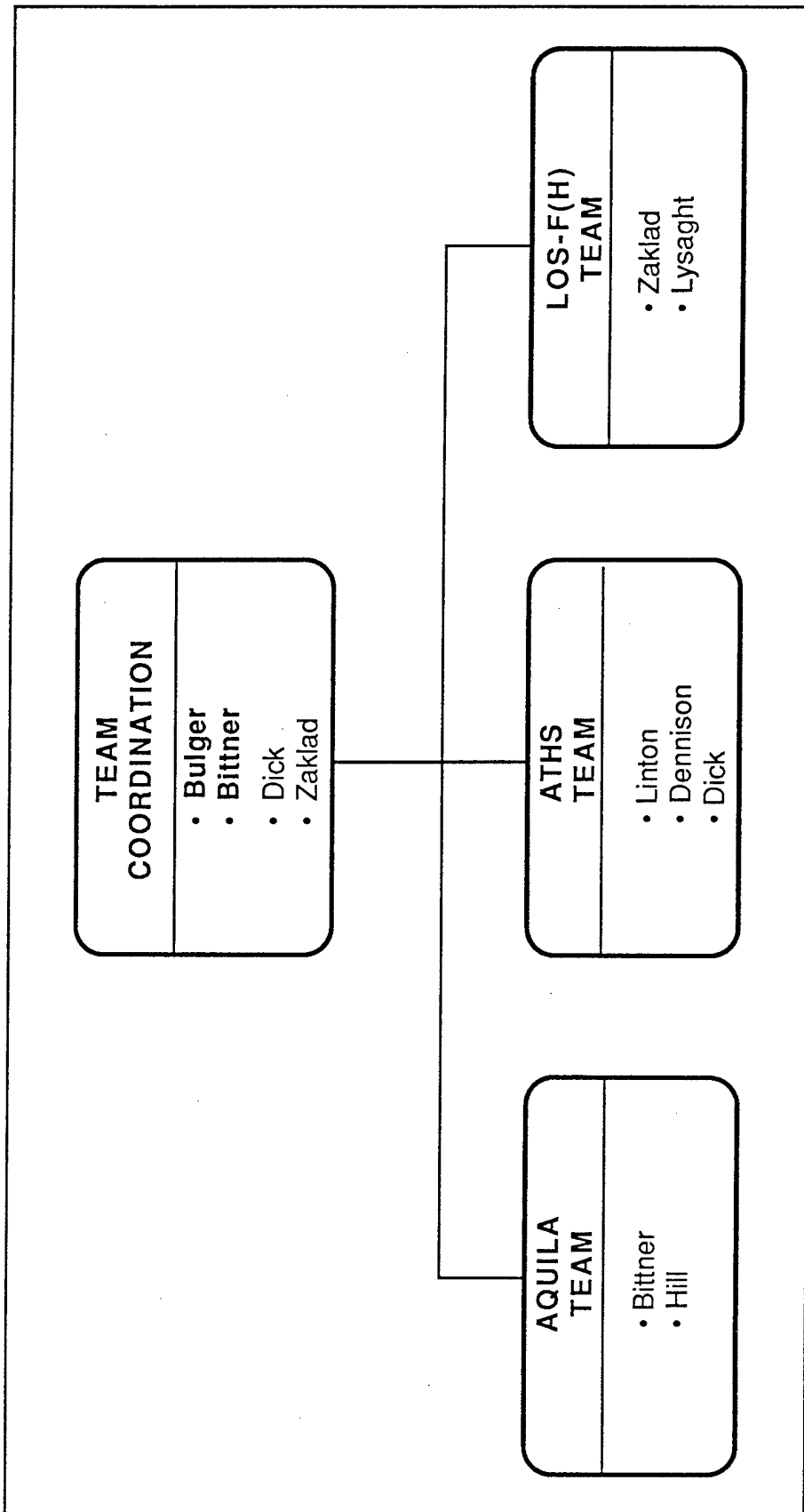
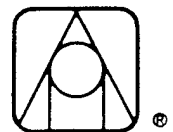


Figure 6.2-1: Validation Staffing Plan

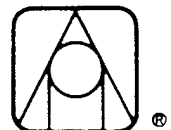


7. REFERENCES

- Adkins, S. R., Levy, L., Muller, D. G., Stinson, D. R. & Rodriguez, L. (1987). *Final report for the Line-of-Sight Forward Heavy cost and training effectiveness study*. US Army Air Defense School (Eds.).
- Allen Corporation of America (1987). *Final report on the preliminary training development study for the Short Range Air Defense System and the Forward Area Air Defense System*. US Army Air Defense School.
- Allen, M. J., & Yen, W. M. (1979) *Introduction to measurement theory*. Belmont, CA: Wadsworth, Inc.
- Anderson, T.W. (1984). *An introduction to multivariate statistical analysis (2nd ed.)*. New York: Wiley.
- Bittner, A. C., Jr., Carter, R. C., Kennedy, R. S., Harbeson, M. M., & Krause, M. (1986). Performance Evaluation Tests for Environmental Research (PETER): Evaluation of 114 measures. *Perceptual and Motor Skills*, 63, 683-708.
- Bittner, A. C., Jr., Christ, R. E., Wierwille, W. W., Linton, P. M., & Glenn III, F. A. (1987). Operator workload (OWL) program for Army systems: Overview and progress. In *Proceedings of the Human Factors Society 31st Annual Meeting* (in press). Santa Monica, CA: Human Factors Society.
- British Aerospace/United Technologies (1987). *Forward Area Air Defense Weapon System Line of Sight Component, Tracked Rapier Hybrid, Executive Summary*.
- Butler, D. D., Dylla, F. T., Hulsey, W. F. (1987). *Combined operational/technical Nondevelopmental Item (NDI) candidate evaluation of the Line-of-Sight Forward Heavy (LOS-F(H)), System Test Design Plan*. US Army Air Defense Board.
- Carter, R. C., & Bittner, A. C., Jr. (1982). *Jackknife for variance analysis of multifactor experiments* (NBDL 82R013). New Orleans, LA: Naval Biodynamics Laboratory. (NTIS AD-A121760).
- Chronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.
- Damos, D. L. (1984). Individual differences in multiple-task performance and subjective estimates of workload. *Perceptual and Motor Skills*, 59, 567-580.
- Damos, D. L. (1985) The relation between the type A behavior pattern, pacing, and subjective workload under single- and dual-task conditions. *Human Factors*, 27, 675-680.
- Dunlap, W. P., Silver, N. C., & Bittner, A. C., Jr. (1986). Estimating reliability with small samples: Increased precision with averaged correlations. *Human Factors*, 28, 685-690.



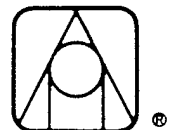
- Dynamics Research Corporation (1983). *Final Report on the Application of the HARDMAN Methodology to the Army Remotely Piloted Vehicle (RPV)*. Technical Report No. R-408U.
- Dynamics Research Corporation (1985). *Reexamination of Support Requirements of the Remotely Piloted Vehicle (RPV) System* (Draft). Technical Report No. E-10053U.
- Eggleston, R. G. (1984). A comparison of projected and measured workload ratings using the Subjective Workload Assessment Technique (SWAT). In *Proceedings of the National Aerospace and Electronics Conference* (pp. 827-831). Dayton, OH: IEEE.
- Foss, C. S. (1987). FAAD Rivals Promote Equipment. *James Defense Weekly* (p. 1281).
- Glenn, F. A., Dick, A. O., & Bittner, A. C., Jr. (1987). Prediction of personnel requirements for system operation. In *Proceedings of the Human Factors Society 31st Annual Meeting* (in press). Santa Monica, CA: Human Factors Society.
- Harris, R. M., Iavecchia, H. P., Ross, L. V., & Shaffer, S. C. (1987). Microcomputer Human Operator Simulator (HOS IV). In *Proceedings of the Human Factors Society 31st Annual Meeting* (in press). Santa Monica, CA: Human Factors Society.
- Hart, S. G., Childress, M. E., & Hauser, J. R. (1982). Individual definitions of the term "workload." In *Proceedings of the Eighth Symposium on Psychology in the Department of Defense* (pp. 478-485). Colorado Springs, CO: USAFA.
- Hershberger, M. L., Murray, S. A., Fitzhenry, P. and Farnochi, A. J. (1983). *Evaluation of Army Piloted Vehicle Mission Payload Operator Performance in Simulated Artillery Missions* Hughes Aircraft Co., Display Systems Laboratory Fort Monmouth, N.J.: US Army Electronics R&D Command (DELCS-TR-81-0133-F).
- Hill, S. G., Lysaght, R. J., Bittner, A. C., Jr., Bulger, J., Plamondon, B. D., Linton, P. M., & Dick, A. O. (1987). *Operator workload (OWL) assessment program for the Army: Results from requirements document reviews and user interview analysis* (TR 2075-2). Willow Grove, PA: Analytics, Inc.
- Hill, S. G., Plamondon, B. D., Wierwille, W. W., Lysaght, R. J., Dick, A. O., & Bittner, A. C., Jr. (1987). Analytic techniques for the assessment of operator workload. In *Proceedings of the Human Factors Society 31st Annual Meeting*. Santa Monica, CA: Human Factors Society.
- Hinkley, D.V. (1983). Jackknife methods. In: S. Kotz, N.L. Johnson, & C.B. Read (Eds.). *Encyclopedia of Statistical Sciences, Vol. 4*, (pp. 280-287). New York: Wiley.
- Horizons Technology Inc. (1986). *Line of Sight - Forward (Heavy) Surrogate Assessment*. Fort Bliss TX: US Army Research Institute for Behavioral Science .



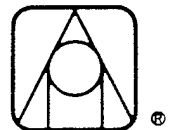
- Horizons Technology Inc. (1986). *Line of Sight - Forward (Heavy) Systems Manprint Management Plan (SMMP). Candidate Evaluation and Selection*. US Army Research Institute for Behavioral Science: Fort Bliss TX .
- HQ DA (1981). *Draft Operators Manual, Guided Missile Intercept - Aerial: AN/GSG-11(V)1 (TM 9-1425-625-10)* .
- HQ DA *Army Training and Evaluation Program Remotely Piloted Vehicle Battery (ARTEP)6-517* .
- HQ DA (1986). *Technical Operators Manual for Remotely Piloted Vehicle (RPV) Target Acquisitions/Designation and Aerial Reconnaissance System (TADARS)* (DEP55-1550-200-10). Volumes 1-5. Lockheed Missile and Space Company.
- Hughes Aircraft Co./Missile Systems Group (1987). *Forward Area Air Defense (FAAD) Line of Site (LOS) - Forward Heavy, Roland.*
- Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in the randomized block and split-block designs. *J. Educ. Statist.*, 1, 69-82.
- Huynh, H., & Feldt, L. S. (1980). Performance of traditional F tests in repeated measures designs under variance heterogeneity. *Commun. Statist. Theory. Meth.* A9(1), 61-74.
- Krohn, G.S. (1987). MANPRINT Findings for the US Army Remotely Piloted Vehicle (RPV). *Appendix B of the USAOTEA Draft Test Report* Essex Corp. Human Factors and Training Group, Westlake Village, CA.
- HQ DA (1986). *Technical Manual Operators for Remotely Piloted Vehicle (RPV) Target Acquisitions/Designation and Aerial Reconnaissance System (TADARS)* (DEP55-1550-200-10) Volumes 1-5 Lockheed Missile and Space Company.
- Kuperman, G. G. (1985). Pro-SWAT applied to advanced helicopter crewstation concepts. In *Proceedings of the Human Factors Society 29th Annual Meeting* (pp. 398-402). Santa Monica, CA: Human Factors Society.
- Lane, N., Strieb, M. I., & Leyland, W. E. (1979). Modeling the human operator: Applications to system cost effectiveness. (NATO/AGARD Proceedings No. CP-268). In *Modeling and Simulation of Avionics Systems and Command Control and Communication Systems*.
- Lysaght, R. J., Hill, S. G., Dick, A. O., Plamondon, B.D., Linton, P. L., Wierwille, W. W., Zaklad, A. L., Bittner, A. C. Jr. (1987). *Operator workload assessment program for the Army: Comprehensive review and evaluation of operator workload methodologies* (TR 2075-3A). Willow Grove, PA: Analytics, Inc.
- Oerlikon-Buhrie LTD (1987). *ADATS Air Defense Anti-tank Systems*. Zurich: Switzerland.



- Reid, G. B., Eggemeier, F., & Nygren, T. (1982). An individual differences approach to SWAT scale development. In *Proceedings of the Human Factors Society 26th Annual Meeting* (pp. 639-642). Santa Monica, CA: Human Factors Society.
- Reid, G. B., Shingledecker, C. A., Hockenberger, R., & Quinn, T. J. (1984). A projective application of the subjective workload assessment techniques. In *Proceedings of the National Aerospace and Electronics Conference* (pp. 824-826). Dayton, OH: IEEE.
- Scheffe, H. (1959). *The Analysis of variance*. (Chapter 10). New York: Wiley.
- Shifrin, C. A. (1987). Army Tests Line-of-Sight Air Defense Systems. *Aviation Week & Space Technology* .
- Unknown (1978). *Proposed RPV Platoon Composition, MOS's and ASI* .
- US Army Air Defense School (1987). *Evaluation Plan for Line-of-Sight Forward Heavy, Force Development Test and Experimentation, Phase I* .
- US Army Field Artillery School (1986). *Field Circular - Remotely Piloted Vehicle (Aquila) Operations* (Draft).
- US Army Field Artillery School (1985). *Target Acquisition Designation and Aerial Reconnaissance System (TADARS) Remotely Piloted Vehicle (RPV). Operational and Organizational Plan (O & O Plan)* .
- US Army Field Artillery School (1987). *System MANPRINT Management Plan (SMMP) for the Remotely Piloted Vehicle (RPV) Target Acquisitions Reconnaissance/Designated System (TADARS)* (Draft).
- US Army Human Engineering Laboratory (1987). *Human Factors Engineering Analysis for Remotely Piloted Vehicle Systems* . Aberdeen Proving Ground, MD.
- US Army Missile Command (1987). *Capstone Test Evaluation Master Plan (TEMP) for the Forward Area Air Defense Systems (FAADS) Program (S)*.
- US Army Missile Command (1987). *FAAD Test Management Handbook* .
- US Army Missile Command (1987). *Line of Sight-Forward - Heavy (LOS-F-H) Weapon System Master Program Plan*.
- US Army Missile Command (1987). *Test and Evaluation Master Plan (TEMP) for the Forward Area Air Defense (FAAD) Line-of-Sight Forward-Heavy (LOS-F-H) (C)*.
- US Army Missile Command (1987). *Test and Evaluation Master Plan (TEMP III) for the Aquila Remotely Piloted Vehicle (RPV)*.
- US Army Training and Doctrine Command (1986). *Capstone Required Operational Capability for the Forward Area Air Defense System (FAAD) (S)*.
- US General Accounting Office (1986). *Aquila Remotely Piloted Vehicle Recent Developments and Alternatives*.



- Wierwille, W. W., Casali, J. G., Connor, S. A., & Rahimi, M. (1985). Evaluation of the sensitivity and intrusion of mental workload estimation techniques. In W. Rorer (Ed.), *Advances in Man-Machine Systems Research* (Volume 2, pp. 51-127). Greenwich, CT: J.A.I. Press.
- Winer, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.
- Winer, B. J. (1971). *Statistical principles in experimental design*. (2nd ed.). New York: McGraw-Hill.
- Zachary, W., Zaklad, A., & Davis, D. (1987),. *A cognitive approach to multisensor correlation in an advanced tactical environment*. Paper presented at 1st Tri-Service Data Fusion Symposium, Johns Hopkins University, Columbia, MD.
- Zaklad, A., Deimler, J., Iavecchia, H., and Stokes, J. (1982). *Multisensor Correlation and TACCO Workload in Representative ASW and ASUW Environment* (TR 1753A). Willow Grove, PA: Analytics.



Bliss

Working Paper

WP FB 88-01

PEDESTAL MOUNTED STINGER CANDIDATE EVALUATION:

OVERVIEW OF MANPRINT DATA

Rene J. dePontbriand, Ph.D.
Fort Bliss Field Unit, and
E. Wayne Frederickson, Ph.D.
Horizons Technology, Inc.



**U.S. Army Research Institute
for the Behavioral and Social Sciences**
5001 Eisenhower Avenue, Alexandria VA 22333

This working paper is an unofficial document intended for limited distribution to obtain comments. The views, opinions, and/or findings contained in this document are those of the author(s) and should not be construed as the official position of ARI or as an official Department of the Army position, policy, or decision, unless so designated by other official documentation.

FB 9039

PEDESTAL MOUNTED STINGER CANDIDATE EVALUATION:
OVERVIEW OF MANPRINT DATA

ABSTRACT

The forward area air defense Pedestal Mounted Stinger (PMS) non-developmental item candidate evaluation was conducted during the period 10 November 1986 through 1 July 1987. Along with other test activities, Manpower and Personnel Integration (MANPRINT) data collection and analyses were carried out. Of the six areas covered under MANPRINT, the two PMS candidates were rated as follows. (1) Human Factors, Soldier Performance: Candidate A ahead. Human Factors, Engineering: No clear choice. (2) System Safety: No clear choice. (3) Health Hazards: No clear choice. (4) Manpower: Candidate A ahead. (5) Personnel: Candidate A ahead. (6) Training: No clear choice. Overall, trends in the available data favored Candidate A over both Candidate B and Man-Portable Air Defense System (MANPADS). The more diagnostic forms of data will be used in further development of the winning system as fielding time approaches.

PEDESTAL MOUNTED STINGER CANDIDATE EVALUATION:
OVERVIEW OF MANPRINT DATA

CONTENTS

	Page
Introduction	1
Background.....	1
System Description.....	3
Caveats.....	4
PMS MANPRINT Evaluations.....	6
Overview of Data.....	6
Human Factors.....	7
Soldier Performance.....	7
Human Factors Engineering.....	23
System Safety and Health Hazards.....	27
Manpower, Personnel and Training.....	28
Manpower.....	28
Personnel.....	30
Training.....	41
Conclusions.....	42
References.....	46
Appendix A: Human Factors Engineering.....	A
Appendix B: System Safety and Health Hazards.....	B
Appendix C: Preliminary Training Requirements Analysis....	C

LIST OF TABLES

Table 1-1. March order times.....	8
1-2. Emplacement times.....	9
1-3a. One missile reload time.....	10
1-3b. Eight missiles reload time.....	10
1-3c. Gun reload time.....	10
1-4. Conversion times.....	12

CONTENTS (Continued)

1-5. Engagement outcomes by system, all targets.....	14
1-6. MANPRINT crew performance errors, all targets.....	14
1-7a. MANPRINT detection errors.....	15
1-7b. MANPRINT tracking errors.....	15
1-7c. MANPRINT firing errors.....	17
1-8a. Nonrecoverable MANPRINT errors, rotary wing trials, POST.....	18
1-8b. Nonrecoverable MANPRINT errors, fixed wing trials, POST.....	21
1-9a. MANPRINT gunner performance errors, CanA, POST.....	24
1-9b. MANPRINT gunner performance errors, CanB, POST.....	25
1-9c. Individual crewman performance summary, POST.....	26
2-1. Individual crewman characteristics and performance.....	32
2-2. PMS task list (JASS analyses).....	36
2-3. Human abilities (Fleishman taxonomy).....	39
2-4. Operator aptitude weights and rankings.....	40
3-1. Training questionnaire summary.....	43

LIST OF FIGURES

Figure 1. MANPRINT team, Pedestal Mounted Stinger non-developmental item candidate evaluation....	2
---	---

(COMPETITION SENSITIVE)

PEDESTAL MOUNTED STINGER CANDIDATE EVALUATION:
OVERVIEW OF MANPRINT DATA

INTRODUCTION

Background

The forward area air defense (FAAD) Pedestal Mounted Stinger (PMS) non-developmental item candidate evaluation (NDICE) was conducted during the period 10 November 1986 through 1 July 1987 at Oro Grande, New Mexico. In addition to conducting normal test activities, the PMS NDICE provided an opportunity to evaluate Manpower and Personnel Integration (MANPRINT) concerns in an operational test setting.

This working paper (WP) is intended to present an overview description of results from that MANPRINT evaluation, and to provide information to be used in the PMS System MANPRINT Management Plan (SMMP) process. To better orient the reader to the processes these data were put through, (a) evaluative comments are provided in those sections for which there was greatest concern within the test and evaluation community, and (b) where most appropriate, the discussion combines data from several MANPRINT areas when that provides a more complete description of an identified problem.

MANPRINT-related participants in the PMS NDICE are listed in Figure 1. Overall MANPRINT coordination was provided by the Army Research Institute (ARI), Systems Research Laboratory (SRL), Fort Bliss Field Unit. The MANPRINT portions of the Test Design Plan, data collection, data reduction, and data analyses were carried out by ARI, its MANPRINT support contractor Horizons Technology (HTI), and the Army Materiel Test and Evaluation (ARMTE) Directorate.

On 12 February 1986, the Directorate of Combat Developments, Air Defense Artillery School, requested ARI assistance in detailing MANPRINT test plans for the FAAD PMS (formerly known as line of sight-rear). ARI provided Stinger troop performance data previously obtained with ARI's Realistic Air Defense Engagement System (RADES). On 24 September 1986, the Deputy Chief of Staff for Personnel, requested that ARI get directly involved in the upcoming FAAD tests. The first such test was the PMS NDICE; ARI was requested by the Air Defense Artillery Board (hereafter referred to as "Board") and PMS NDICE Test Manager, to provide MANPRINT support for the entire period covered by this evaluation. The senior author acted in the role of PMS NDICE MANPRINT technical coordinator and team leader.

The purpose of the PMS NDICE was to conduct an

PMS MANPRINT TEAM

Team Leader: Dr. dePontbriand, ARI

Support Staff: Mr. Moran, Army Materiel Test and Evaluation (ARMTE)

Dr. Frederickson, Horizons Technology, Inc (HTI)

Mr. Balarzs, Human Engineering Laboratory (HEL)

Mr. Medina, ARMTE

Field and Data Collection Staff Leaders:

Mr. Welch, Combat Systems Test Activity (CSTA); vehicle and predicted fire weapon.

Mr. Franks, Army Environmental Hygiene Agency (AEHA); radiation, LASER.

Mr. Moran, ARMTE; human factors engineering, toxic gas, noise.

Mr. Frydendall, Mr. Brown, (HTI); soldier performance, manpower, personnel and training.

Data Analysis Group Staff Leaders:

Mr. Doyle, HTI

Mr. Ramirez, Air Defense Artillery Board

**Figure 1: MANPRINT Team, Pedestal Mounted Stinger
Non-developmental Item Candidate Evaluation.**

(COMPETITION SENSITIVE)

operational and technical test of the candidate systems; results were to be provided to the Source Selection Board. The Operational Test and Evaluation Agency (OTEA) was the operational independent evaluator; the Board was the operational tester. The Army Materiel Systems Analysis Activity (AMSAA) was the technical independent evaluator; ARMTE was designated the technical tester.

The NDICE was conducted in two phases, in accordance with the government's PMS request for proposals (RFP). In Phase I, Candidate A (CanA), Candidate B (CanB) and the man-portable air defense system (MANPADS) were tested side by side. Operating simultaneously, to help ensure that all tested systems would engage targets under similar circumstances, these systems attempted to detect, identify, track and simulate engagement of aircraft flying in accordance with the threat support package. Tactical movement, operational time and expected threat were based on the FAAD operational and organizational (O&O) plan. Systems were tested for acquisition and tracking using (a) Stinger Basic Trainer (Basic) and (b) Stinger passive optical seeker technology (POST) configurations against single aircraft. Basic trials made up the early portion of Phase I only; the majority of data reported herein concern POST trials.

Tests were conducted to assess candidates' survivability and detectability, between weapons-systems electromagnetic compatibility, rapid successive engagement capability (demonstration only), laser range finder performance, self-defense reaction time, tracking quality, missile tipoff. To the extent possible, candidates also were evaluated for MANPRINT concerns, reliability, availability, and maintainability (RAM), built-in test/built-in test equipment (BIT/BITE), and integrated logistics support (ILS).

In Phase II, which immediately followed Phase I, candidate systems underwent testing in environmental chambers in terms of temperature extremes and varying rain conditions. BIT/BITE concerns were addressed further at that time.

System Description

The following two paragraphs are taken verbatim from the Board's PMS NDICE report, dated 27 October 1987.

"System Background. The need for a forward area air defense (FAAD) capability has been analyzed by the Army FAAD Working Group. This analysis resulted in the identification of a family of air defense component systems to be fielded in the FAAD Battalion of maneuver divisions. A more detailed discussion of the need for this component system can be found in the FAAD Program Capstone Test and Evaluation Master Plan

(COMPETITION SENSITIVE)

(TEMP) and the FAAD Working Group Final Report. The PMS is the NDI solution to the line-of-sight-rear (LOS-R) component of the FAAD program. The Required Operational Capability (ROC) document for the PMS system was approved on 20 May 1986. The O&O plan was approved on 6 June 1986. As a component of the FAAD program, the PMS is organic to the FAAD battalion and will be employed in the rear battle area of the heavy, light, and special divisions, Armored Cavalry Regiments (ACR), and corps air defense brigades. The PMS is a mobile system designed to provide counter air protection against low-altitude, fixed- and rotary-wing threat aircraft in all operational environments. The PMS provides low-altitude air defense protection for both critical stationary and mobile assets in support of the force commander's concept of the operation.

"System Description. The FAAD concept integrates systems designed to counter the low-altitude air threat over and beyond a division's area of operations. This program consists of line-of-sight (LOS) weapons and non-line-of-sight (NLOS) weapons; command, control, and intelligence (C2I) sensors (both ground and aerial, using passive and active technologies); aircraft identification techniques (active friendly and passive hostile identification); automated command and control (command post, integration, and distribution of data among FAAD elements and to other forces, and interface with Joint Air Defense C2I); and the required communications to support information distribution. The PMS system is expected to be the NDI solution for LOS-R component requirements in the FAAD program."

The basic PMS configuration includes two four-Stinger pods and a predicted fire weapon (machine gun), all mounted on a high-mobility, multi-wheeled vehicle (HMMWV), and with a 360-degree azimuth slew and -10 to +70 degree elevation capability. Also included is a forward-looking infrared radar (FLIR). The PMS has capability for remote operations to a range of 50 meters from the base system. Other operational and employment concept descriptions can be found in the Board's full report. CanA had a driver/observer in the HMMWV cab; the gunner was positioned in the turret, which had a plexi-glass canopy cover. CanB had both driver/observer and gunner positioned in the HMMWV cab. The MANPADS team was deployed as a two-man team using a HMMWV for mobility. The designated military occupational specialty (MOS) was 16S, MANPADS operator.

Caveats

1- This report contains information retrieved from: (a) the Board's report, cited above, (b) the Board's NDICE data archives, and (c) preliminary MANPRINT results and reports. Certain segments of the test data are classified

(COMPETITION SENSITIVE)

"Confidential"; these will be released only by the Board or in OTEA and AMSAA Independent Evaluation Reports. Reports submitted by ARMTE, addressing human factors engineering, system safety and health hazards are summarized in this paper, and included in their entirety in the appendix to this report.

2- Due to the nature of data reported, this document is classified "Competition Sensitive".

3- Due to the classification level, distribution of this WP is limited and will be determined by ARI SRL.

4- Based on a 28 September 1987 telephone conversation with COL Drolet, Stinger Program Manager, the data in this report are releasable under the following conditions: (a) Candidates are not to be identified by name; (b) results must be coded, e.g., Candidate A, Candidate B.

5- Raw data and computer printouts are available on a limited access basis from the Board; POCs, COL Pedigo, AV 978-3000, LTC Moss, 978-1030.

6- Due to the broad nature of the MANPRINT data, some of the troop performance results were not labeled specifically as MANPRINT. However, those results were generated through or coordinated with the MANPRINT portion of the project. Those data and the results in question were instead labeled as and included under "Performance" in reports to the Source Selection Evaluation Board (SSEB) and OTEA, the two primary recipients of all test data. Integration of all operator data was carried out in meetings with these two agencies. Such an integration is consistent with ARI's understanding of "System Performance" as including a soldier performance component as well as equipment (government, contractor) performance components.

7- All data reported in this report are based on single-target trials. The limited number of multiple-target trials conducted are described in the Board's report.

8- MANPADS was included in this candidate evaluation (a) to provide an empirical baseline against which to compare PMS candidate performance and (b) to demonstrate under similar test conditions whether the PMS can overcome MANPADS shortcomings due to very limited nighttime capability.

9- Due to the small subject sample available, only descriptive and data are presented.

10- In an effort to facilitate candidate comparisons, performance differences are reported in the text generally in terms of percentages. Since some of the numbers used in computing these percentages are quite small in absolute terms,

(COMPETITION SENSITIVE)

reference to the data tables provided is advisable.

PMS NDICE MANPRINT EVALUATIONS

Overview of Data

The critical soldier performance data which are classified "Confidential" and not reported here, come under the following general data categories:

- Launch times
- Identification-to-launch times
- Acquisition-to-launch times
- Launch range for engageable hostile targets
- Identification range for hostile targets
- Handoff times for engageable hostile targets
- Acquisition range for hostile targets
- Detection range for hostile targets

Soldier performance scores are reported for the following conditions:

- March order, emplacement, reload and conversion to MANPADS
- Fixed wing (FW) and rotary wing (RW) aircraft targets
- Cued and uncued targets
- Mission oriented protective posture (MOPP) 0 and MOPP IV
- Benign and infrared counter measures (IRCM)
- Day and night
- Stationary, moving, and remote engagements

The following sections deal with the six domains of MANPRINT:

1- Human factors (HF), in terms of human factors performance, which is the contribution of soldier performance to system performance; and in terms of human factors engineering (HFE), which considers such factors as anthropometrics and adequacy of controls and displays. The essential HF concern was stated as, "Can PMS operators employ the system successfully under potential operating conditions?".

2- Health hazards (HH), which looks principally at the effects of environmental or operational conditions from the perspective of chronic, long-term hazard to the soldier. The essential concern was: "Are PMS operators adequately protected against health hazards in the operational environment?"

3- System safety (SS), which looks principally at dangerous conditions from the perspective of acute, immediate hazards to the soldier. The essential concern was: "Can the PMS system be operated safely under potential operating

(COMPETITION SENSITIVE)

conditions?"

4- Manpower, which looks at the adequacy of planned crew size in terms of completing all necessary tasks in a given time frame. This area also examines the Army force structure's ability to provide sufficient personnel to properly field the system, a concern only partially within the scope of the immediate test. The essential concern was: "Is the proposed 2-person crew suitable or adequate for PMS operations?"

5- Personnel, which looks at the skills and abilities available in the projected military occupational specialty (MOS) to meet the weapon's skills and abilities requirements. The essential concern was: "Do PMS operations impose personnel requirements (e.g., Armed Services Vocational Aptitude Battery (ASVAB), physical profile) not available in the 16S pool?"

6- Training, which looks at the ability of the Army's training support structure to meet training requirements within the training resources currently needed for the transition MOS. The essential concern was: "Will PMS require training that cannot be provided within the available/planned 16S training base?"

Human Factors

Soldier Performance

Soldier performance data were obtained for several distinct activities. Time trials for march order, emplacement, reload and MANPADS conversion (conversion of the PMS to the MANPADS configuration) were conducted by ARMTE and ARI/HTI support personnel. Troops followed procedures as outlined in their operator's manuals. While there are between-candidate systems differences, all times were considered to be within acceptable limits by the Stinger-PM and SSEB (unless otherwise noted).

March order. PMS candidate march order times are listed in Table 1-1. In all cases, CanA had shorter absolute median times for remote conditions, while CanB had shorter absolute median times for stationary conditions.

Emplacement. Emplacement times are listed in Table 1-2. In those cases where CanB was quicker than CanA (again all stationary), there was only a small absolute time difference; in those cases where CanA was quicker (again all remote), there was a relatively higher absolute time savings.

Reload. Reload times are found in Tables 1-3a to 1-3c, and include times for single-missile reload, eight-missile

(COMPETITION SENSITIVE)

Table 1-1: March Order Times

<u>Condition</u>	<u>Candidate</u>	<u>Sample Size</u>	<u>Median (sec)</u>	<u>Maximum (sec)</u>	<u>Minimum (sec)</u>
Stationary	CanA	20	99	180	49
	CanB	16	37	75	26
Remote	CanA	16	131	230	79
	CanB	17	414	737	269
Remote Day	CanA	13	128	230	79
	CanB	13	414	660	269
Remote Night	CanA	3	155	208	112
	CanB	4	448	737	336
Stationary Day	CanA	13	80	145	49
	CanB	10	37	75	26
Stationary Night	CanA	7	91	180	76
	CanB	6	44	56	32
Remote MOPP 0	CanA	15	131	230	79
	CanB	16	403	737	269
Remote MOPP IV	CanA	1	128	128	128
	CanB	1	414	414	414
Stationary MOPP 0	CanA	15	101	145	69
	CanB	12	37	75	26
Stationary MOPP IV	CanA	5	99	180	49
	CanB	4	37	56	32

(COMPETITION SENSITIVE)

Table 1-2: Emplacement Times

<u>Condition</u>	<u>Candidate</u>	<u>Sample Size</u>	<u>Median (sec)</u>	<u>Maximum (sec)</u>	<u>Minimum (sec)</u>
Stationary	CanA	17	108	211	72
	CanB	15	96	371	37
Remote	CanA	20	254	431	131
	CanB	17	443	631	270
Remote Day	CanA	15	255	363	131
	CanB	13	449	631	336
Remote Night	CanA	5	254	431	199
	CanB	4	394	613	270
Stationary Day	CanA	11	108	211	80
	CanB	9	98	371	37
Stationary Night	CanA	6	107	166	72
	CanB	6	75	97	55
Remote MOPP 0	CanA	14	251	431	131
	CanB	13	395	631	270
Remote MOPP IV	CanA	6	288	427	201
	CanB	4	508	613	443
Stationary MOPP 0	CanA	13	108	211	72
	CanB	11	97	371	37
Stationary MOPP IV	CanA	4	105	166	92
	CanB	4	85	115	46

(COMPETITION SENSITIVE)

Table 1-3a: One Missile Reload

<u>Condition</u>	<u>Candidate</u>	<u>Sample Size</u>	<u>Median (sec)</u>	<u>Maximum (sec)</u>	<u>Minimum (sec)</u>
Day	CanA	23	48.0	121	26
	CanB	21	30.0	46	19
Night	CanA	17	49.0	121	33
	CanB	19	29.0	60	21
MOPP 0	CanA	28	49.0	121	30
	CanB	28	30.5	60	19
MOPP IV	CanA	12	41.5	80	26
	CanB	12	28.5	46	20

Faster times in MOPP IV may be attributed to the fact that the majority of MOPP IV trials were conducted at end of Phase I when soldiers had more experience.

Table 1-3b: Eight Missiles Reload

<u>Condition</u>	<u>Candidate</u>	<u>Sample Size</u>	<u>Median (sec)</u>	<u>Maximum (sec)</u>	<u>Minimum (sec)</u>
Day MOPP 0	CanA	5	170	290	161
	CanB	5	106	108	80

Table 1-3c: Gun Reload

<u>Condition</u>	<u>Candidate</u>	<u>Sample Size</u>	<u>Median (sec)</u>	<u>Maximum (sec)</u>	<u>Minimum (sec)</u>
Day MOPP 0	CanA	4	450.5	496	335
	CanB	4	408.5	463	261

(COMPETITION SENSITIVE)

(full complement) reload, and gun reload. CanB consistently took less time for reloads across all conditions.

MANPADS conversion. Based on ROC and RFP requirements, any missile on the PMS must be capable of conversion to MANPADS by representative soldiers within certain time limits (limits to be determined at a later date by TRADOC). Conversion times for day, night, MOPP 0, MOPP IV and over all conditions are presented in Table 1-4. CanB crew conversion time was consistently shorter than CanA conversion time.

Operability. The following engagement sequence operability data were collected and analyzed, and came from (a) real-time and post-mission test-site observations and debriefings and (b) gunner-station audio/video tape data reduction. The analytical focus was on detection of crew performance errors and determination of error causes. Brief definitions are as follows:

1- MANPRINT error: An operator engagement-sequence performance error (faulty or incorrect actions; omission of action; actions performed out of proper sequence).

2- MANPRINT error rate: Total number of MANPRINT errors for a given condition divided by the total number of valid trials for that condition (these data dealt only with non-recoverable errors, hostile targets).

3- Valid trials: All trials conducted in which government equipment (e.g., target aircraft and communications) was fully operational and performing in accordance with test criteria.

4- Recoverable MANPRINT error: An operator error which does not result in the premature termination (prior to missile launch or gun firing) of an engagement. Several recoverable errors could be committed in either successful or unsuccessful engagements.

5- Nonrecoverable MANPRINT error: An operator error which resulted in the premature termination of the engagement sequence. This was also referred to as "engagement ending error". A nonrecoverable error resulted in an unsuccessful engagement.

6- Successful engagement: A trial in which a validly presented hostile target was fired upon such that intercept would have occurred within the engagement boundary of the missile or gun (either weapon, as appropriate).

7- Unsuccessful engagement: A trial in which a validly presented hostile target was not fired upon with either the

(COMPETITION SENSITIVE)

Table 1-4: Conversion Times

<u>Condition</u>	<u>Candidate</u>	<u>Sample Size</u>	<u>Median (sec)</u>	<u>Maximum (sec)</u>	<u>Minimum (sec)</u>
All	CanA	41	83.0	136	40
	CanB	44	53.5	126	25
Day	CanA	22	87.0	120	40
	CanB	24	52.5	112	25
Night	CanA	19	74.0	136	56
	CanB	20	64.5	126	44
MOPP 0	CanA	29	87.0	136	56
	CanB	32	51.5	126	25
MOPP IV	CanA	12	67.0	100	40
	CanB	12	62.0	104	44

Note: The PMS system must be capable of conversion to the MANPADS configuration in case of HMMWV or or power subsystem malfunction. All times reported here were judged acceptable by the Stinger-PM.

(COMPETITION SENSITIVE)

missile or gun within their respective engagement boundaries; a validly presented hostile target was fired upon outside the engagement boundaries; or, a friendly target was fired upon.

Overall system operability for CanA, CanB and MANPADS is compared for Basic and POST trials in Table 1-5. For all comparisons shown in Table 1-5, CanA was better than CanB. Both candidates performed better than MANPADS on POST trials; both candidates demonstrated better performance during POST trials than during basic trials.

MANPRINT Crew Performance Errors. This analysis of the operational engagement sequence was based on the commonly accepted event sequence shown in Table 1-6. Table 1-6 indicates where in this sequence MANPRINT errors occurred and their frequency of occurrence. Two points will be noted. The first of these deals with the proportion of engagement ending errors to overall errors for each system. For CanA, engagement ending errors make up 64 percent of overall errors; for CanB, the proportion is 46 percent. However, the second point is more critical to overall performance. This deals with the overall number of engagement ending errors--those errors which lead to an unsuccessful mission or missed target opportunity due to operator errors. Here, CanB had 13 percent more errors than did CanA (computed as $[\text{number of CanB errors}] - [\text{number of CanA errors}] / [\text{number of CanA errors}]$). For both candidates, the events with the highest frequency of engagement ending errors were detect, tracking, and fire. These three error loci were analyzed further to determine more specific MANPRINT-related causes.

Causes of MANPRINT Errors. As can be seen in Table 1-6 and detailed below, the highest incidence of nonrecoverable errors occurred during detection, tracking and firing. Probable causes of the more frequent classes of errors were obtained during structured interviews with PMS crews and review of the crew-station video.

For detection errors (Table 1-7a), the most frequent problem cited for both candidates was seeing the target. Other problems cited dealt with a mix of procedural, training, and human factors engineering concerns. Both candidates had difficulty handing off the target or search sectors between observer and gunner; CanB crews had nearly twice as many detection errors due to handoff problems as did CanA. CanB also experienced errors due to use of wrong procedures and detection of the wrong target.

Performance in the tracking phase (Table 1-7b) appeared to be affected by the time spent trying to detect the target. Crews were still trying to line up the target when the "Cease Engagement" order was given from the command post, an action

(COMPETITION SENSITIVE)

Table 1-5: Engagement Outcomes By System for All Targets

<u>System</u>	<u>Successful Trials (%)</u>		<u>Unsuccessful Trials (%)</u>	
	<u>Basic</u>	<u>POST</u>	<u>Basic</u>	<u>POST</u>
CanA	60	80	40	20
CanB	50	73	50	27
MANPADS*	59	59	41	41

*All Basic trials for MANPADS were daylight trials; all other trials data reported represent a day/night mix.

Table 1-6: MANPRINT Crew Performance Errors
All Targets (POST)

<u>Event</u>	<u>CanA</u>		<u>CanB</u>	
	<u>Recoverable Error</u>	<u>Engagement Ending Error</u>	<u>Recoverable Error</u>	<u>Engagement Ending Error</u>
Detect	6	49	8	74
Acquire	0	4	1	6
Identification	19	9	24	5
IFF	33	0	135	0
Tracking	13	33	0	30
Ranging	0	0	0	0
Fire	0	34	0	31
Slew to Cue	<u>1</u>	<u>0</u>	<u>4</u>	<u>0</u>
Total Errors	72	129	172	146

(COMPETITION SENSITIVE)

Table 1-7a: MANPRINT Detection Errors

	CanA	CanB
Never Saw Target	35	41
Handoff Problem	14	27
Wrong Search Procedures	0	4
Detected Wrong Target	<u>0</u>	<u>2</u>
Total Detection Errors	49	74

Table 1-7b: MANPRINT Tracking Errors

	CanA	CanB
Late Detection/Contact	8	20
Tracked Cloud or Flare	7	4
Thought Target Was Out of Range	7	2
Tracked - No Seeker Lock	2	1
Tracked - No Target ID	5	0
Tracked Wrong Target	4	0
Lost Track at FOV Change	<u>0</u>	<u>3</u>
Total Tracking Errors	33	30

(COMPETITION SENSITIVE)

which signified that the target aircraft was no longer in the engagement boundaries. The most frequently cited cause for tracking errors was "late detection or contact", with CanB showing 150 percent more errors than CanA (CanB, 20; CanA, 8). However, in all but one other cause category, CanB outperformed CanA. CanA crews mistakenly tracked a cloud or decoy flare more often; lost track because they thought the target was out of range; tracked but did not identify the target; and tracked the wrong target. CanB lost track at the field of view (FOV) change three times. This loss appeared to be attributable to the suddenness with which the system would go from wide to narrow FOV, which made it difficult to keep the target within view, and to a requirement to momentarily move the hand from the tracking thumbwheel control to activate the FOV control. CanA's FOV change was done in a "zoom-in" manner, allowing the target to be kept in the center of the screen, and did not require moving the hands from controls.

As for identified causes of firing errors (Table 1-7c), CanA crews tended to fire out of bounds when there was no system range information available, that is, when they were called upon to do manual target ranging. Compared to CanB, CanA committed 280 percent more errors in this category. Both systems tended occasionally to fire out of bounds even with the availability of range information; such problems may be candidates for correction through use of automated firing lock-out functions.

Nonrecoverable MANPRINT errors, RW trials, POST. Non-recoverable errors were further analyzed to determine the effect of various test conditions on crew performance. Results for hostile RW targets are presented in Table 1-8a. Total error rate for each candidate was 11 percent, compared to 39 percent for MANPADS. To a large degree, the poor MANPADS performance was attributable to its low nighttime capabilities. Cuing level had no measurable effect for CanA; CanB's performance improved in the uncued mode. Level of MOPP had no effect on performance in RW trials. Some of these results appeared due to the fact that the slow-moving RW targets provided an inordinate amount of time for engagement; creeping along the desert floor at the test site, these targets could sometimes be seen several minutes before cuing was provided by the command post.

Mode of operation (stationary, moving, remote) did have an affect on the performance of each system. In stationary, systems engaged from an emplaced position. In moving, systems engaged while moving over pre-determined trails, each system crossing over similar terrain features of gravel, small hills, and light brush. In remote, soldiers parked the vehicle upon receiving orders from the command post, and deployed to a

(COMPETITION SENSITIVE)

Table 1-7c: MANPRINT Firing Errors

	CanA	CanB
Fired Out-of-Bounds No Range Information	19	5
Fired Out-of-Bounds With Range Information	7	9
Fired in Missile Dead zone	2	6
Fired Without Seeker Lock	6	7
Fired at Flare	0	3
Failed to Arm System	<u>0</u>	<u>1</u>
Total Firing Errors	34	31

(COMPETITION SENSITIVE)

Table 1-8a: Nonrecoverable MANPRINT errors,
Rotary Wing trials, POST.

<u>Test Condition</u>	CanA		CanB		<u>MANPADS</u>	
	<u>No. Passes</u>	<u>Error Rate (%)</u>	<u>No. Passes</u>	<u>Error Rate (%)</u>	<u>No. Passes</u>	<u>Error Rate (%)</u>
Benign	177	11	185	11	188	39
IRCM	0					
Cued	130	11	141	13	136	37
Uncued	47	11	44	5	52	42
MOPP 0	141	11	151	11	154	38
MOPP IV	36	11	34	12	34	41
Stationary	50	4	51	8	47	36
Moving	73	12	72	15	67	35
Remote	54	15	62	8	74	43
Day	130	5	114	11	110	6
Night	<u>47</u>	<u>26</u>	<u>71</u>	<u>10</u>	<u>78</u>	<u>85</u>
Total	177	11	185	11	188	39

(COMPETITION SENSITIVE)

remote position up to 50 meters distant which was chosen by the crew to optimize their perspective of the designated search sector. The MANPADS HMMWV followed along the same trails; stationary and remote were essentially similar operations, while for moving, the crew would stop upon receiving the cue and attempt to engage the target.

CanA crews exhibited an error rate of 4 percent in the stationary mode; this increased by 3 to 4 times in moving and remote. For CanB, the error rate in moving was twice as high as in either stationary or remote. CanA outperformed CanB in stationary and moving.

The remoting capability performance was of particular concern in this test, based in part on O&O concerns and the different ways the two candidates implemented the feature (CanA had two separate fire control consoles (FCC); CanB unbolted the single FCC from the main system). Crewmen for both candidates reported difficulties with the FCC; however, CanA's error rate (15 percent) was nearly twice that of CanB's (8 percent).

CanA crewmen complained about the remote FCC prototype's weight of 65 pounds; projections are for a finished weight of approximately 35 pounds. Crewmen were observed failing to reintegrate the remote FCC with the main system's turret on several occasions, after first integrating the remote unit properly but then moving it again. Troops reported that the remote unit was moved because either the initial positioning made for very uncomfortable gunner seating; or, upon having set up the unit, the gunner found he was not provided an adequate view of the designated search sector. Thus, remote site selection and integration routines presented training and procedures concerns, both of which can be viewed as at least partially correctable. A different kind of problem with the remote unit was the differing location of controls, switches and displays from their relative location on the main system FCC. See Appendix B, Human Factors Engineering, for a broader discussion.

It should be noted here that one crewman's performance accounted for 48 percent of overall engagement sequence operator errors on CanA. That individual's profile included corrective lenses, left handedness, and other anomalies compared to the other test troops. While such individual differences will be discussed more completely in the personnel section, they are mentioned here because they appear pertinent to the discussion above.

CanB crewmen complained about the difficulty in handling and connecting cables to the remote unit. Cables had to be passed through small holes in the vehicle cab, and crewmen

(COMPETITION SENSITIVE)

would thereby scrape their knuckles. Despite repeated difficulties in handling the unit, a difficulty which showed up in relatively poor emplacement and march order times (reported above), CanB's remote engagement performance was better than CanA's. In addition to the personnel concern cited above for CanA, the better remote performance for CanB in part was seen as due to the fact that CanB used the same FCC in all conditions. This naturally avoided problems with adjusting to a different man-machine interface, a problem which appeared to negatively affect performance on CanA.

In the day condition, CanA's error rate was slightly less than half that of CanB's (5 percent vs 11 percent), and similar to the error rate for MANPRINT (6 percent). In the night condition, CanA's error rate was 2.6 times higher than CanB's (26 percent vs 10 percent). CanA's relatively poor night performance appeared, again, to be due in part to the performance of the same individual cited above.

Nonrecoverable MANPRINT errors, FW trials, POST. Results for hostile FW targets are presented in Table 1-8b. CanA had the best performance overall with its 21 percent error rate, compared to CanB's 26 percent and the MANPADS' 42 percent. Compared to RW trials, CanA's error rate was 91 percent higher; CanB's was 136 percent higher; MANPADS' error rate was higher by 9 percent. Across all conditions, the only time CanB performed better than CanA was in the remote condition, for reasons cited under the previous RW discussion. Except for the anticipated day/night performance difference, MANPADS showed little performance variability either between conditions or within levels of the other conditions.

CanA outperformed CanB in the conditions of IRCM, by 40 percent; uncued, by 24 percent; MOPP IV, by 79 percent. CanA also was best in stationary, by 200 percent; moving, by 81 percent; and night, by 57 percent.

IRCM had its most negative effect on CanB. CanB soldiers would mistakenly track flares for too long and not recover in time to successfully engage the true target. Results in this condition were likely affected by the CanB contractor decision to change the FLIR subsystem several times during the conduct of the test. While these changes were intended to improve overall system performance, it could be speculated that they may also have resulted in some degree of negative transfer, (e.g., in timing, rhythm, "signature") since each of the three FLIRs had its own peculiar operating characteristics. However, in the absence of any further detailed information, it could also be speculated that this poor performance was due to hardware problems.

In the uncued condition, both candidates had more

(COMPETITION SENSITIVE)

Table 1-8b: Nonrecoverable MANPRINT errors,
Fixed Wing trials, POST.

<u>Test Condition</u>	CanA		CanB		<u>MANPADS</u>	
	<u>No. Passes</u>	<u>Error Rate (%)</u>	<u>No. Passes</u>	<u>Error Rate (%)</u>	<u>No. Passes</u>	<u>Error Rate (%)</u>
Benign	347	20	326	24	331	44
IRCM	91	25	94	35	99	34
Cued	316	20	298	24	319	42
Uncued	122	25	122	31	111	41
MOPP 0	353	21	329	22	342	42
MOPP IV	85	24	91	43	88	42
Stationary	129	8	132	24	136	40
Moving	165	21	144	38	141	43
Remote	144	34	144	17	153	42
Day	278	21	273	21	284	16
Night	<u>160</u>	<u>23</u>	<u>147</u>	<u>36</u>	<u>146</u>	<u>91</u>
Total	438	21	420	26	430	42

*This relatively favorable IRCM performance for MANPADS was due to their lack of availability, and hence reliance on, a FLIR. If the MANPADS seeker locked onto a dropped flare, the operator could see the error and go back in search of the target.

(COMPETITION SENSITIVE)

difficulty against FW than against RW targets; again, this appeared attributable to the difference in time the two types of targets were available for engagement. CanA's better performance in this condition may be attributed in part to its turret-based configuration, which provided for better viewing of the search sector by the gunner than was available for the crew cab-enclosed gunner compartment for CanB.

This same time availability difference between RW and FW targets may have contributed to the increase in MOPP IV error-rate over the RW target trials. CanA showed a 118 percent increased error rate against FW than against RW targets, while CanB experienced a 291 percent increase. Within the FW trials, CanA's MOPP IV error rate was 14 percent higher than for MOPP 0; for CanB, the error rate was 95 percent higher in MOPP IV than in MOPP 0. Between candidates, the MOPP IV superiority for CanA (by 79 percent) appeared to be due to both controls/displays layout and amount of room available in the turret compared to CanB's vehicle-cab crew compartment.

There was no effect of MOPP level for MANPADS. While overall performance diminished for both CanA and CanB in FW vs RW trials, MANPADS performance showed only moderate differences between and comparatively poor performance in these two types of trials. Except for obvious day-night differences, and the effect of IRCM, the MANPADS error rate was relatively constant.

In the operational mode conditions, CanA performed substantially better in both stationary and moving modes, but worse in the remote mode.

A recurring problem with CanB had to do with the manual tracking mechanism, a thumb wheel. That mechanism was one of the factors which tended to negatively affect stationary performance (24 percent error rate); such factors were even more of a problem in the moving mode (38 percent error rate). CanA's two-handed function-balanced controls provided added stability, which contributed to better performance during stationary and moving modes.

CanA performance in the remote mode again was affected negatively, at least in part due to problems cited above with gunner A4 and the differences in operator consoles; CanA's error rate in remote was twice as high as CanB's.

CanA performance showed little effect of night operations, showing a 9.5 percent increase over day; CanB performance indicated a high negative influence, with a 71.4 percent increase in error rate over day trials. It is not clear why these results are so different from the parallel condition in RW trials.

(COMPETITION SENSITIVE)

As was expected, MANPADS showed its worst performance at night (91 percent error rate), with a 469 percent increase over day trials (16 percent error rate). This was one place where the PMS concept showed its advantage over MANPADS. The other anticipated strong point favoring PMS over MANPADS was in the area of multiple engagements, since the PMS is designed with eight available missiles at any one time. However, except for a brief proof-of-principle demonstration, not reported upon here, all trial data in this report were collected under single-target conditions.

MANPRINT gunner performance by engagement event, non-recoverable errors, POST. Some members of the evaluation community were concerned only with the "performance" aspects of MANPRINT. In response to their interest in the various probable causes of performance differences, this section reviews the number of engagement ending errors committed by individual gunners for events in the engagement sequence. Only valid hostile targets were used in this analysis. CanA data are presented in Table 1-9a, CanB data in Table 1-9b. These data will be revisited in the personnel section.

Gunner A3 was the best performer for CanA, and in fact was the best gunner over both candidates. Gunner A4 was the poorest performer overall, and was substantially poorer for all but one event that was associated with any errors. Gunners A1 and A2 performed about the same; their performance was better than the performance of any of the CanB gunners.

Summary performance figures by gunner are presented in Table 1-9c. One finding here is the obvious difference in error rates among CanA gunners, compared to fairly consistent performance among CanB gunners.

Human Factors Engineering

The summary information in this section relates to compliance with MANPRINT regulations and human factors engineering standards as defined in AR 602-1, AR 602-2, and MIL-STD-1472, and was collected by ARMTE with partial support from ARI/HTI personnel. See Appendix A for the ARMTE report.

CanA concerns. There are many differences in the arrangements of controls, switches and indicators between the gunner's FCC on the system and the FCC used in remote operations. For example, some switch placements were reversed, and some controls were missing, on the remote FCC. Differences in these consoles appeared to most directly affect operator performance in remote. There were other inconsistencies and noncompliance with the labeling on these consoles. In all, there were 25 citations of noncompliance with regulations.

(COMPETITION SENSITIVE)

Table 1-9a: MANPRINT Gunner Performance Errors
CanA Gunners (POST)

<u>Event</u>	<u>A1</u>	<u>A2</u>	<u>A3</u>	<u>A4</u>	<u>Total</u>
Detect	11 22%	13 27%	1 2%	24 49%	49 38%
Acquire	0	0	0	4 100%	4 3%
Identification	1 11%	1 11%	1 11%	6 67%	9 7%
IFF	0	0	0	0	0
Tracking	9 27%	5 15%	2 6%	17 52%	33 26%
Ranging	0	0	0	0	0
Fire	9 26%	10 29%	4 12%	11 32%	34 26%
Slew to Cue	0	0	0	0	0
Overall	30 23%	29 22%	8 6%	62 48%	129 100%

(COMPETITION SENSITIVE)

Table 1-9b: MANPRINT Gunner Performance Errors
CanB Gunners (POST)

<u>Event</u>	<u>B1</u>	<u>B2</u>	<u>B3</u>	<u>Total</u>
Detect	22 30%	32 43%	20 27%	74 51%
Acquire	2 33%	0	4 67%	6 4%
Identification	2 40%	2 40%	1 20%	5 3%
IFF	0	0	0	0
Tracking	6 20%	12 40%	12 40%	30 21%
Ranging	0	0	0	0
Fire	12 39%	6 19%	13 42%	31 21%
Slew to Cue	0	0	0	0
Overall	44 30%	52 36%	50 34%	146 100%

(COMPETITION SENSITIVE)

Table 1-9c: Individual Crewman Performance Summary (POST)

<u>Crewman</u>	<u>Percent Passes Successful</u>		<u>MANPRINT Error Rate</u>	
CanA				
A1	80		17	
A2	84		14	
A3	92		5	
A4	66		32	
CanB				
B1	68		21	
B2	74		20	
B3	76		18	
<u>MANPADS</u>				
	<u>All</u>	<u>Day</u>	<u>All</u>	<u>Day</u>
D1	63	81	32	13
D2	56	89	42	9
D3	63	80	34	14
D4	53	85	43	10

NOTE: Percent successful trials and MANPRINT error rate do not total 100 percent because some unsuccessful trials are due to other than MANPRINT causes.

(COMPETITION SENSITIVE)

CanB concerns. CanB also exhibited noncompliance with control console labeling, although the incidence is somewhat higher for CanA. For all categories of noncompliance, including labeling, there were 18 separate citations.

System Safety and Health Hazards

The information in this section relates to PMS candidates conformance to criteria established in Army regulations documents. The principal system safety (SS) and health hazards (HH) assessments effort was conducted by ARMTE; ARI support personnel observations, incidents reports and other assistance were utilized where needed. The full text of the ARMTE assessment is reported in Appendix C. Findings are summarized here.

Carbon monoxide and nitrogen dioxide. For both CanA and CanB, levels of measured vehicle exhaust gasses at the crew stations were well below Occupational Safety and Health Administration (OSHA) limits for carbon monoxide (CO) and nitrogen dioxide (NO₂).

Hydrogen chloride. For hydrogen chloride (HCL) levels, acceptable limits have not yet been established. There are currently two standards. The OSHA standard is 7mg/m³ or 5 ppm; the National Institute of Occupational Safety and Health (NIOSH) standard is listed as 150 mg/m³ or 100 ppm. For CanA, no HCL was detected in the gunner's (turret) station during this test (for more extensive measurement levels in the vehicle cab, please see appendix, Table E-7). After earlier test firings in which missile backblast forced open the cab doors, improved blast and gas protection was provided; however, as time progressed after these firings, it appears that the presence of HCL in the cab worsened even with the improvements, but were still generally within the less stringent NIOSH standards. The structural improvements did, however, keep the doors from being forced open; remaining problems were due to the door seals.

For CanB, the same effect of backblast occurred, this time severely buckling the system's doors. It took two modifications, blast deflectors and additional cab door latches, to repair this problem. HCL levels were generally within NIOSH standards.

Based on standards described in MIL-STD-882B, in the worst case when the missile backblast is directed at the cab, the high HCL levels entering the cab are classified as "catastrophic" in severity and "probable" in frequency for both CanA and CanB. CanA's gunner compartment was well protected.

(COMPETITION SENSITIVE)

Environmental temperature. For both candidates, the hazard of heat stress is classified as "critical" in severity; neither candidate system came with an air conditioner.

Impulse noise. Based on the 140 decibel (dB) contour measurements, it was determined that hearing protection would be needed within 575 feet of the CanA system during missile firing; the 140 dB contour for CanB was 519 feet. In both cases, the noise levels were higher for missile firings than for predicted fire weapon (PFW) firings; for both systems, the PFW 140 dB contour was 288 feet. Conditions would call for crew-cab personnel to wear double hearing protection (e.g., ear plugs and helmet) during missile firings (this holds only for when the cab doors are blasted open, as occurred during the NDICE; further evaluations were scheduled for the winning candidate in early December 1987). Single hearing protection (ear plugs) only would be required during PFW firings.

SS and HFE. The following SS findings resulted from human factors engineering evaluations. For CanA, an entryway-located radio antenna was often mistakenly used as a handhold by gunners; this antenna (a) is too flexible to serve as a stable handhold and (b) can give a radio frequency (RF) burn if gripped while the radio is transmitting. This hazard is classified as "critical" in severity and "probable" in frequency, but could be controlled by use of warning labels or, preferably, by relocation. Several "cutting" (sharp edge or corner) hazards are classified as "marginal" in severity and "occasional" in frequency. The most critical safety and health concern related to system performance during live firings. Again, further evaluations are scheduled.

For CanB, it is possible for the system to fire a missile into the radio antennas, thereby damaging both communications and the missile. This hazard is classified as "critical" in severity and "remote" in frequency. HMMWV filter capacitors can be caused to discharge and cause a spark during normal operations of battery replacement; the battery damage is classified as "marginal" in severity and "remote" in frequency.

Manpower, Personnel and Training

ARI conducted all phases of the manpower, personnel and training (MPT) analyses. There were several limitations in the scope of these analyses; major among these was the restricted sample size of available troops. Another critical limitation to the scope of MPT issues in general was that candidate-specific maintainer requirements could not be adequately addressed; all maintenance was provided by candidate contractor engineers. Other limitations are noted where they were considered to have the greatest impact.

(COMPETITION SENSITIVE)

Manpower

As stated earlier, the PMS NDICE manpower issue was the suitability or adequacy of size of the designated two-man crew. The manpower issue would generally include force structure concerns, such as whether the Army had enough 16S MOS personnel to field PMS. While this latter consideration was outside the scope of this test, it was indirectly addressed, though not fully answered, with results from the personnel analysis. That analysis suggests that 16S personnel who require corrective lenses may not be appropriate for inclusion as PMS crewmen; to the extent that this finding holds up under further evaluation, affected individuals would have to be removed from the ranks of available soldiers. Clearly, personnel and manpower issues are interrelated.

The primary approach to addressing adequacy of crew size was an empirical, operationally oriented procedure used to assess operator performance, or system operability. Performance was evaluated in terms of errors and event times. If an engagement was not successfully completed, the event during which the mission ending error was committed was identified and the cause diagnosed as accurately as possible. The data used for assessing operator performance consisted of the MANPRINT computerized database, crewstation and through-sight audio/video tapes, data collection forms from field observations, and post-mission crew debriefing forms.

Essentially, data generated for the soldier performance aspect of human factors were utilized to identify any potential crew-size deficiencies. The engagement sequence task event which would benefit most from added crew members would be detection, the most error-prone event for either candidate (see Table 1-6). However, revised procedures should be considered first; such changes are the purview of the Force Development Test and Experimentation (FDT&E) process. System design changes should also be considered. Additionally, performance times for march order, emplacement and reload might well benefit from increased crew size, but NDICE-derived performance times for those tasks were considered by the Stinger-PM office and the SSEB to be within limits of acceptability.

An operator workload (OWL) analysis was initiated. This analysis was derived from the McCracken and Aldrich (1984) workload measure developed for the Army's Light-Helicopter Experimental (LHX) program, chosen because of the high level of cognitive tasks contained in both LHX and PMS operations. For the PMS, with a period of time counted in seconds, integrating and acting upon information from a C2I source, from a crewmember, from one's own search of the designated search sector, and attempting to identify threat aircraft on

(COMPETITION SENSITIVE)

the FLIR, ranging, etc., all contributed to the analyst's judgment of a high cognitive component requirement for system operation. However, failures to obtain reliable task time data resulted in a workload measure which could only be interpreted as a measure of "perceived difficulty" of the tasks, not workload as it is generally understood. For this and related reasons, workload data are not reported here. The results are included in the Board report, and are available strictly for archival purposes. The less than satisfactory experience encountered in the attempt to directly assess OWL in this study does demonstrate the need for improved operator workload measurement techniques, ideally those which would be unobtrusive to ongoing operational testing and transparent to the participant.

The third and final form of manpower-related data was provided by the Man-Integrated Systems Technology (MIST) analysis (Stewart & Shvern, 1987). This was the closest this MANPRINT effort came to considering maintainer requirements. The MIST PMS maintainer analysis used data from the PMS O&O plan, the ROC document, and the Light Air Defense System (LADS) hardware and manpower (HARDMAN) analysis. These MIST data are a projection of the maintenance manpower needs for a generic PMS system, and do not apply specifically to a particular candidate system.

Briefly stated, the MIST findings strongly suggest that BIT/BITE estimates currently entertained are optimistic, and that maintainer workload will be significantly higher than anticipated. The recommendation from the analysis is that more of the maintainer task burden be reallocated from the anticipated maintainer (24X MOS) to some new operator/maintainer (16X MOS). However, according to the authors of that report, the types of tasks which could most directly be shifted off to the 16X are those least likely to significantly reduce the 24X maintenance burden. (As of this writing, the notion of "operator/maintainer" is at best a back-burner issue.) While the generic MIST information did not directly bear upon the PMS NDICE decision, it will be necessary to track the issue through the acquisition process, and thus enter it into the PMS SMMP process.

Personnel

It warrants repeating that the most challenging aspect of a personnel or any MANPRINT analysis in an NDICE of this type is the small sample of subjects available. In light of that constraint, these data should at best be utilized as guidance for the planned FDT&E. It also can be speculated that since a large number of trials were conducted across several months and across varying conditions, results from these data should provide some confidence in the findings. It behooves the user

(COMPETITION SENSITIVE)

community to make the best use of the best data available.

As was true for manpower analyses, personnel concerns addressed in this study could not look explicitly at personnel requirements for maintainers; it is anticipated that citing such experiences will provide a lesson learned to readers in the MANPRINT community. As stated above, maintenance was provided by contractors, whose performance was not directly evaluated; such performance data would have been of only moderate utility in any case, since they would not have been derived from a representative sample of prospective PMS maintainers. RAM figures, included in the full Board report, may provide relevant information as to levels of maintainer workload requirements based in the documented profile of equipment reliability.

Two specific, if limited, analyses were conducted to examine the personnel issue. These analyses first looked at the relationship between each soldier's Operator/Foods (OF) aptitude area (AA) score and General Technical (GT) AA score and their individual test performance. OF and GT scores were obtained through TRADOC System Manager (TSM) offices at the ADA School. The OF score is used as the criterion score for selection to the 16S MOS. GT is a score that is highly correlated with the Armed Forces Qualification Test (AFQT) score, which is seen as a measure of general learning ability. (The GT scores were the designated "surrogate" for AFQT scores.) Table 2-1 presents a comparison between OF and GT scores in relation to two performance measures: (a) percent of successful passes, and (b) MANPRINT error rate, closely related to measure (a). These two scores do not total to one-hundred percent since some unsuccessful trials were due to other than MANPRINT type errors. Of the two, MANPRINT error rate represents a more purely MANPRINT measure.

There is evidence that CanB may be affected by an undetermined system limitation which restricts even those with high aptitudes to performance at the same level as those with lower aptitudes; such a condition is not evident in CanA data. Based on the very limited data available, there is a stronger relationship between OF scores and percentage of successful trials for the CanA crew than for the CanB crew. There was a range of 26 percentage points for the "successful passes" performance scores for CanA, with soldier A3 scoring a 92 and soldier A4 scoring 66. In line with this, there was a 22 point spread in the OF scores for CanA, with A3 scoring highest and A4 lowest. The range of performance for CanB was 8 percentage points, reflecting more similarity of performance across CanB crewmen, even though the OF score range was 32 points. The same relative results obtain in the comparison of the "error rate" scores and OF.

(COMPETITION SENSITIVE)

Table 2-1: Individual Crewman Characteristics and Performance

<u>Crewman</u>	<u>OF Score</u>	<u>GT Score</u>	<u>Percent Passes Successful *</u>		<u>Error Rate Due to MANPRINT Causes</u>		
			<u>All</u>	<u>Day</u>	<u>All</u>	<u>Day</u>	
<u>CanA</u>							
A1	106	107	80		18		
A2	102	89	84		14		
A3	119	110	92		5		
A4	97	97	66		32		
<u>CanB</u>							
B1	92	87	68		21		
B2	124	117	74		20		
B3	118	112	76		18		
<u>MANPADS</u>							
			<u>All</u>	<u>Day</u>	<u>All</u>	<u>Day</u>	
D1	98	108	63	81	32	13	
D2	115	107	56	89	42	9	
D3	93	88	63	80	34	14	
D4	127	124	53	85	43	10	

*Both MANPRINT and non-MANPRINT errors taken into account.

(COMPETITION SENSITIVE)

During SSEB discussions, a question arose regarding whether CanA operators would benefit from raising OF score requirements from 90 to 100. Based on the limited analyses possible here (indicating that the one below-100 OF scorer was the worst performer), such appears to be the case. More importantly, the overall AA-score issue requires further analysis in light of recent proposals to redefine the criteria for admission to the 16-series MOS.

A note must be made in reference to individual differences, a core personnel concern which is critical to a full understanding of these test data. Although much effort was put into selection to ensure representativeness of test soldiers, the small selection pool restricted both selection and placement of individuals. In the best light, it must be said that all soldiers had just completed Advanced Individual Training (AIT), a factor which aids in ensuring at least some degree of homogeneity among individuals to be distributed across candidate systems. However, after extensive contractor and on-site training, coaching, practice and feedback as well as other forms of "special" treatment, it is not clear just how representative of the target audience test participants remained once the test got underway. But whatever representativeness may have been lost due to such treatment, it would have been a fairly equal loss across candidates.

With regard to soldier A4, there appeared to be a potentially anomalous situation which could unrealistically and negatively skew CanA system performance scores. The following information was compiled on 4 August 1987 and provided to OTEA and to the SSEB.

1- In ratings provided by test soldiers upon their return to Ft. Bliss from contractor-provided training (see training section), A4 is a consistent negative outlier in his perceptions of training in "Time to Learn", "Learning Difficulty", "Performance Difficulty", and "Error Probability" in 12 task areas; A4's ratings were close to teammates' ratings only in "Importance to Mission" of the twelve task areas.

2- A4 was fairly representative of 16Ss on selection and profile criteria, except for two potentially salient features. The first of these is that he is left-handed, the only left-handed soldier assigned to either candidate. An analysis of the CanA controls indicates a balance of functions between right and left hands, with many switches operable by either hand. It is clearly not as handedness-critical as CanB, which has the manual tracking thumbwheel mounted on a joystick for right-hand operation. The second differentiating feature is that A4 wears corrective lenses; as with handedness, he was the only test participant with this differentiating feature.

(COMPETITION SENSITIVE)

There exists a variety of reasons why corrective lenses might negatively affect performance on a PMS-type system. For example, many wearers of corrective lenses are either near-sighted or far-sighted; corrective lenses, at least in the form of eyeglasses vs contact lenses, adjust for the faulty condition at the expense of impairing the healthy condition. That is, by correcting near-sightedness, eyeglasses often render the wearer farsighted. In PMS operations, it is necessary to be able to read both display and control information at arm's length and search the designated target sector involving distances in the range of 10 kilometers. Other potential problems arise due to the several layers of light refraction affecting the eyeglass wearer: corrective lens, laser protection lens, MOPP IV mask, and windshield or canopy. Other reasons may account for A4's performance, but the corrective lens concern cannot be answered satisfactorily with NDICE data alone.

3- In a post-test question and answer session at the PMS SSEB reviews, CanA troops were asked about the effect on visibility of rain and/or dust on the turret canopy. In an informal estimation of how much this rain/dust situation degraded visibility, A1, A2 and A3 estimated perhaps as much as 30-50 percent degradation of visibility resulted. However, even after clarification questioning by the Test Officer, A4's stated estimate was 70 percent degradation. Again, whether due to eyeglasses or some other factor cannot be determined, and the issue remains to be resolved.

It was the position of MANPRINT personnel, and one supported by OTEA, that A4's performance did not necessarily reflect an accurate picture of what would occur in a normal Unit setting. There, such anomalous behaviors which could negatively affect overall crew performance or morale would trigger some corrective action. In a Unit setting, someone with A4's profile might be teamed with someone more appropriate to optimizing system performance, might receive one-on-one tutoring, or receive some form of within-Unit counseling. In the extreme, such an individual might even leave the situation due to peer or other pressure. Whatever the remedy, it is presumed that performance, attitudes, and behaviors exhibited during the test and reflected in A4's performance scores would be rectified to some extent in an actual Unit situation.

As the PMS test scoring situation existed, CanA's overall position would change from somewhat better than CanB's to much better than CanB's, on the active performance measures cited in this report.

The second form of personnel analysis involved the

(COMPETITION SENSITIVE)

application of ARI's Job Assessment System Software (JASS) methodology. JASS helps identify skills and aptitudes essential to performing system tasks. Test players provided input to the JASS program for the determination of requisite skill aptitudes for each system. This database consisted of player ratings of aptitude requirements for a set of selected tasks for their respective systems (see Table 2-2). The aptitudes rated are derived from Fleishman's taxonomy of perceptual, cognitive and psychomotor abilities (Table 2-3). Table 2-4 lists the rated aptitude weights and ranks for MANPADS, CanA and CanB.

Ratings were obtained after completion of the PMS NDICE, by the only trained and experienced personnel available, the test soldiers themselves. PMS NDICE soldiers provided the ratings. Their common basis of reference was that they had all been trained on MANPADS, as part of AIT training, and operated MANPADS as part of the "MANPADS Conversion" exercise which constituted part of the test. Based on the data presented in Table 2-4, the MANPADS-CanA aptitude requirements rank order correlation was .76; MANPADS-CanB correlation was .58. On the basis of these limited data, which precluded statistical analyses, there appears to be a closer match of aptitude requirements between MANPADS, the feeder system for PMS operators, and CanA, than appears to exist between MANPADS and CanB.

This correlational data suggests that the personnel (and relatedly, training) systems would experience less disruption, or enjoy a smoother transition, in going from MANPADS to CanA operations than from MANPADS to CanB operations. More of the aptitudes required (and available) in the target audience of 16S would be put to use in operation of CanA than in operation of CanB. Such data tend to support performance scores cited in this report as favoring CanA.

Since one intended use of this report is the support of the PMS SMMP process of tracking MANPRINT issues through system fielding, and since some readers will look only at sections of particular interest to them, one point will be repeated here. As was the case for several other analyses, while operator performance scores reflect the results of hundreds of trials, comparisons are based upon a very small sample. Such is generally the case for this type of candidate evaluation, wherein soldier support is difficult to secure for the extended length of time required. It is intended that these personnel findings be further explored during the PMS FDT&E Phase II, scheduled for January 1989. At that time, there will be six two-person teams of PMS players. Also at that time, the topic of aptitude requirements can be revisited, as can the issue concerning whether OF score requirements should be raised from 90 to 100 to help ensure

(COMPETITION SENSITIVE)

Table 2-2: CanA PMS Task List

PRACTICE TASK-DRIVING A CAR
EMPLACE FIRE UNIT AT BRIEFED LOCATION
SET-UP SYSTEM FOR STATIONARY OPERATIONS
VISUALLY SEARCH SECTOR
SEARCH SECTOR WITH FLIR
RESPOND TO C2 CUEING OF POTENTIAL TARGETS
DETECT TARGET
ACQUIRE TARGET
MANUALLY TRACK TARGET
IDENTIFY TARGET USING IFF
DETERMINE TARGET RANGE VISUALLY USING MANUAL
RANGEFINDER
ENGAGE TARGET WITH MISSILES
ASSESS TARGET DAMAGE

(COMPETITION SENSITIVE)

Table 2-2 (Cont'd): CanB PMS Task List

PRACTICE TASK-DRIVING A CAR
EMPLACE FIRE UNIT AT BRIEFED LOCATION
SET-UP SYSTEM FOR STATIONARY OPERATIONS
VISUALLY SEARCH SECTOR
SEARCH SECTOR WITH FLIR
RESPOND TO C2 CUEING OF POTENTIAL TARGETS
DETECT TARGET
ACQUIRE TARGET
MANUALLY TRACK TARGET
IDENTIFY TARGET USING IFF
DETERMINE TARGET RANGE VISUALLY USING MANUAL
RANGEFINDER
DETERMINE TARGET RANGE USING LASER
ENGAGE TARGET WITH MISSILE
ASSESS TARGET DAMAGE

(COMPETITION SENSITIVE)

Table 2-2 (cont'd): MANPADS Task List

PRACTICE TASK-DRIVING A CAR
EMPLACE FIRE UNIT AT BRIEFED LOCATION
SET-UP SYSTEM FOR STATIONARY OPERATIONS
VISUALLY SEARCH SECTOR
RESPOND TO C2 CUEING OF POTENTIAL TARGETS
DETECT TARGET
ACQUIRE TARGET
MANUALLY TRACK TARGET
IDENTIFY TARGET USING IFF
DETERMINE TARGET RANGE VISUALLY USING MANUAL
RANGEFINDER
ENGAGE TARGET WITH MISSILE
ASSESS TARGET DAMAGE

(COMPETITION SENSITIVE)

Table 2-3: Human Abilities*

- | | |
|----------------------------|-----------------------------|
| 1. Oral Comprehension | 21. Visualization |
| 2. Written Comprehension | 22. Static Strength |
| 3. Oral Expression | 23. Explosive Strength |
| 4. Written Expression | 24. Dynamic Strength |
| 5. Memorization | 25. Trunk Strength |
| 6. Problem Sensitivity | 26. Stamina |
| 7. Originality | 27. Extent Flexibility |
| 8. Inductive Reasoning | 28. Dynamic Flexibility |
| 9. Category Flexibility | 29. Gross Body Equilibrium |
| 10. Deductive Reasoning | 30. Speed of Limb Movement |
| 11. Information Ordering | 31. Gross Body Coordination |
| 12. Math Reasoning | 32. Multi-Limb Coordination |
| 13. Number Facility | 33. Wrist-Finger Speed |
| 14. Fluency of Ideas | 34. Finger Dexterity |
| 15. Time Sharing | 35. Manual Dexterity |
| 16. Flexibility of Closure | 36. Arm/Hand Steadiness |
| 17. Speed of Closure | 37. Control Precision |
| 18. Selective Attention | 38. Rate Control |
| 19. Perceptual Speed | 39. Reaction Time |
| 20. Spatial Orientation | 40. Choice Reaction |

* Fleishman, E.A. On the Relation Between Abilities and Human Performance American Psychologist, 1975, 30, 1129-1149.

(COMPETITION SENSITIVE)

Table 2-4: Operator Aptitude Weights and Rankings

MANPADS			CanA			CanB		
APTITUDE NAME	WEIGHT	RANK	APTITUDE NAME	WEIGHT	RANK	APTITUDE NAME	WEIGHT	RANK
1	23.0	16	1	14.6	18	1	19.9	15.0
2	19.4	21	2	13.1	22	2	16.2	21.0
3	11.7	28	3	16.0	16	3	16.3	20.0
4	4.8	33	4	3.8	31	4	10.8	26.0
5	58.6	3	5	45.8	4	5	34.1	7.0
6	51.9	4	6	58.8	1	6	49.8	1.0
7	20.7	18	7	17.0	14	7	24.0	13.0
10	6.0	32	10	9.7	25	10	10.3	27.0
11	8.3	31	11	3.2	32	11	4.3	33.0
12	12.1	27	12	5.9	28	12	12.8	23.0
13	13.3	25	13	3.0	33	13	15.8	22.0
14	23.4	15	14	13.9	20	14	12.2	24.0
15	15.4	23	15	13.5	21	15	19.8	16.0
16	13.4	24	16	11.1	23	16	20.5	14.0
17	20.8	17	17	18.1	13	17	3.7	34.0
18	60.7	2	18	56.1	2	18	38.6	6.0
19	60.8	1	19	52.4	3	19	43.8	2.0
20	50.1	5	20	39.4	5	20	39.4	4.5
21	44.3	6	21	23.3	11	21	24.8	12.0
26	2.4	34	26	1.8	34	26	6.9	32.0
27	25.2	13	27	6.6	27	27	10.9	25.0
28	20.4	19	28	5.3	29	28	9.2	29.0
29	32.2	9	29	7.9	26	29	7.6	30.0
30	32.9	8	30	26.6	10	30	19.4	18.0
31	18.0	22	31	3.6	30	31	9.5	28.0
32	12.8	26	32	26.7	9	32	19.1	19.0
33	30.1	10	33	27.0	8	33	39.4	4.5
34	19.6	20	34	14.1	19	34	28.0	10.0
35	24.3	14	35	19.2	12	35	31.1	9.0
36	39.8	7	36	37.5	6	36	39.8	3.0
37	11.5	29	37	15.6	17	37	25.2	11.0
38	8.4	30	38	10.4	24	38	19.7	17.0
39	27.1	11	39	36.7	7	39	33.5	8.0
40	26.8	12	40	16.5	15	40	7.0	31.0

NOTE: These scores represent scores for MANPADS, CanA and CanB averaged across operators, for aptitudes requirements ratings, based on JASS analysis utilizing Fleishman's Aptitude Taxonomy. "1"= highest or most important; "5"= lowest or least important. "Weight" is a relative measure of how important that aptitude is for operation of these systems (e.g., high weighting means high perceived importance compared to items with a lower weight). All ratings were obtained after completion of PMS operational test, by trained and experienced crewmen. MANPADS-CanA rho= .76; MANPADS-CanB rho= .58.

(COMPETITION SENSITIVE)

optimal system performance (a topic which also has manpower and training implications), since there are fewer 16Ss with a score of 100+ than those with a score of 90+.

Training

The final set of MANPRINT analyses had to do with training. Two issues were addressed: (1) will PMS require training that cannot be provided within the available or planned training base, and (2) what was the perceived adequacy of the contractor training provided?

A training requirements analysis (TRA) was conducted using principles outlined by the TRADOC Systems Approach to Training. Procedures included development of a task analysis, determination of institutional training requirements and determination of unit training requirements. Since portions of the text of the TRA report were inadvertently omitted from the Board's report, the complete TRA is given in Appendix C.

To summarize the TRA, few training requirements differences between candidates were identified, and few additional training requirements beyond MANPADS training requirements were found. For Institutional training, there was a total of 18.5 additional hours of training requirements for either candidate over MANPADS. With an expected 16 course repetitions per year (needed to maintain PMS Force status quo), the total additional training requirement for either candidate over current MANPADS requirements is 296 hours. Instructors are available for duty 1,728 hours per year, making the training burden an additional .17 instructors or no additional instructor requirement (the cutoff would be if more than .5 instructors, or 864 additional hours, would be required). Such a finding may not be expected, since the PMS is intended to be more sophisticated than MANPADS. The case is that PMS troops would not undergo currently-required Redeye training; thus, a task training savings is gained at the beginning.

As for Unit training requirements, commanders will maintain the requirement to train 20 percent of individual critical tasks, and to sustain proficiency in all tasks. It is not yet known what embedded or other training devices will be available, and thus such information could not be considered in these evaluations. Thus, it could not be determined how such plans or training devices would have differentiated between candidates.

Contractor training evaluation was conducted immediately after soldiers returned from that training (however, training was continued after reporting to the test site; this follow-on training required a Stinger-PM decision in interpretation of the RFP, as it was not initially planned). Evaluation was

(COMPETITION SENSITIVE)

conducted using a questionnaire. Participants' responses indicating the time to learn, learning difficulty, performance difficulty, importance of the task, and probability of making an error were obtained for the tasks listed in Table 3-1. Table 3-1 also presents a summary of values assigned to the responses on a five-point scale.

Contractor training provided another test constraint, in that there was no opportunity for MANPRINT test personnel to evaluate that training either on-site or through availability of a written outline of the training conducted.

Some correlation may be found in the responses of soldier A4, who rated time to learn and learning difficulty relatively high compared to teammates' ratings, and A4's performance scores. With the exclusion of A4's training evaluation scores, it appears that ratings by the two candidates' crews may well be attributed to differences in training methods; it is not possible to determine actual training difficulty on this limited analysis. A true analysis of that issue could be based on a different type of analysis than that conducted here, one based on an examination of skill decay at different points during the NDICE.

CONCLUSIONS

As a first attempt at applying a full complement of MANPRINT analyses in a Candidate Evaluation, the nature and amount of test information obtained resulted in a qualified success. Certainly, there were many lessons learned; the most critical will be documented in a companion report.

Perhaps the most serious challenge is dealing with the small sample of troops available in such a test. Reliability measures are difficult to apply, although there was the positive condition in this regard of having data from a large number of trials, and data from several different perspectives (i.e., the different areas of MANPRINT), all brought to bear in addressing operational MANPRINT issues.

Along these lines, it is the intention of the MANPRINT program to provide a means for integrating information across its six areas. An example of such integration was combining personnel information (wearing of corrective lenses, ASVAB scores), human factors engineering information (e.g., differences in system versus remote FCCs), health hazards information (in the form of required MOPP gear and laser eye protection) and training information (e.g., the lack of proper training provided to CanA soldiers for setting up the remote FCC) to help interpret performance differences between systems. Past analyses were in part hampered when it came to a broad consideration of soldier concerns, with the various areas of

(COMPETITION SENSITIVE)

Table 3-1: TRAINING QUESTIONNAIRE SUMMARY

	TIME TO LEARN				LEARNING DIFFICULTY				PERFORMANCE DIFFICULTY				IMPORTANCE TO MISSION				ERROR PROBABILITY			
	A4	A3	A2	A1	A4	A3	A2	A1	A4	A3	A2	A1	A4	A3	A2	A1	A4	A3	A2	A1
CanA																				
EFU LOCAL	3	3	2	2	3	1	1	1	2	1	1	1	5	5	5	5	3	2	1	1
EFU REMOTE	3	2	3	2	3	1	2	1	3	2	2	1	5	5	5	5	2	2	2	1
SURVEILLANC	1	2	2	1	1	1	1	1	1	1	1	1	5	5	5	5	3	1	1	2
ETSC	5	1	2	3	3	1	3	1	3	1	1	1	5	5	5	5	5	2	1	2
ETMC	5	1	1	1	5	1	1	1	3	2	1	1	5	5	5	5	5	1	2	4
ETRC	5	5	3	2	5	1	3	3	3	1	2	3	5	5	5	5	2	2	1	1
DFULE	2	2	1	1	2	1	1	1	2	1	1	1	3	3	3	3	5	2	3	1
DFURE	3	3	3	2	3	1	3	2	3	1	2	2	5	5	5	5	2	2	2	2
L/R MISSILE	3	1	2	2	2	1	2	2	1	3	1	2	5	5	5	5	2	2	3	2
L/R GUN	2	1	2	2	2	1	2	2	1	3	1	2	5	5	5	5	2	2	3	2
CONVERT	1	1	3	1	1	1	2	2	1	2	2	2	5	5	5	5	5	1	1	3
PDSOC	3	1	3	3	2	1	2	3	2	1	1	2	3	5	5	4	5	1	1	3
TOTAL	36	25	26	25	29	13	22	20	25	18	16	19	56	58	63	59	38	19	19	22
CanB																				
EFU LOCAL	b1	*b3	b2		b1	*b3	b2		b1	*b3	b2		b1	*b3	b2		b1	*b3	b2	
EFU REMOTE	1	1	1		1	1	1		1	1	1		1	3	3		2	1	1	
SURVEILLANC	2	2	1		2	2	1		1	1	1		5	5	5		2	1	2	
ETSC	2	5	1		1	2	1		1	1	1		5	1	5		2	1	2	
ETMC	1	1	1		1	2	1		1	1	1		5	5	5		2	1	3	
ETRC	2	1	1		1	1	1		1	1	1		3	5	5		1	1	2	
DFULE	2	2	1		1	1	1		1	1	1		5	5	5		2	1	2	
DFURE	1	2	1		1	1	1		1	1	1		5	5	5		2	1	3	
L/R MISSILE	1	2	1		1	1	1		1	1	1		5	5	5		2	1	2	
L/R GUN	1	1	1		1	1	1		1	1	1		5	5	5		1	1	1	
CONVERT	1	1	1		1	1	1		1	1	1		3	5	3		2	1	2	
PDSOC	1	5	1		1	2	1		1	1	1		3	5	3		2	1	2	
TOTAL	16	24	12		13	15	12		12	12	12		52	54	56		20	12	22	

*: Crewman C4 withdrew from the test early on for personal reasons.
 Legend: EFU Local, Remote: Emplace fire unit. ETSC, ETMC, ETRC: Engage target stationary, mobile, remote. DFULE, DFURE: Displace fire unit from local, remote emplacement. L/R missile, gun: Load, reload. Convert: Convert missiles to MANPADS. PDSOC: Perform daily system operation checks.

(COMPETITION SENSITIVE)

concern left to be assessed independent of each other's findings. There was no potential benefit to be achieved from possible synthesis of these results; in part, the set of analyses discussed here has helped demonstrate that such a situation can change. These data can also be used to help assess the impact of indicated requirements changes in the manpower, personnel, training and system design areas.

Other realized challenges had to do with the sometimes heavy burden on evaluation personnel (60 to 80 hour work-weeks were not uncommon), the lack of fully automated data collection instrumentation, the short time allowed for data reduction, analysis, and interpretation, and the seemingly perpetually evolving schedule of the test itself. Fortunately, such potential roadblocks only served to provide incentive and stimulus for broad-based cooperation among all participants, notably the Board, ARMTE, FAAD-TSM, OTEA and the SSEB. As MANPRINT integrator, ARI's challenge could have been much greater than it was.

The SSEB continually requested that MANPRINT issues be boiled down to an essential "CanA or CanB" decision. Critical issues identified in meetings with OTEA, Stinger-PM and the SSEB are presented below.

Human factors engineering evaluations helped account for some of the findings, such as poor performance for CanA in remote operations and FOV change problems for CanB. Linked with performance data, such diagnostic information identifies where improvements are most needed and can best be made.

Problem areas were identified through system safety and health hazards assessments, with neither candidate truly "winning" over the other in these categories. While these data did not turn out to differentiate between candidates in system comparisons, they will be invaluable in future evaluations during FDT&E and initial operational test and evaluation (IOT&E) processes, and in the December 1987 system safety demonstration on the winning candidate.

Manpower and personnel findings were also shown to relate to system performance, and tended to favor CanA when differences were found. It is anticipated that the MIST results will help direct analyses to critical issues during the PMS Logistics Demonstration, scheduled for June 1988. Training analysis results did not differentiate between conditions.

Companion reports to be published can help the reader fill gaps which exist in this short summary of findings. The Board archives contain volumes of information which could only be touched upon in this report; responsible representatives of the MANPRINT community can be granted access to those archives.

(COMPETITION SENSITIVE)

Based on the overall set of MANPRINT analyses reported here, there is a trend favoring CanA. This trend extends through the performance data and, to an extent, the RAM data as reported to the SSEB and included in the Board's test report. The most objective MANPRINT findings dealt with system operability, whereby performance on a defined series of structured events was diagnosed, quantified, and related to overall system performance (see, in particular, Tables 1-6 and 1-7a-c). That is the kind of data which can be linked most directly to strategies to improve battlefield effectiveness of the overall system through improved and more reliable soldier performance.

It is the intention of the PMS MANPRINT community that results from the PMS NDICE contribute to a foundation database for the support of continued FAAD system development through the PMS SMMP, the FDT&E process, and the Pre-Planned Product Improvement process, to the eventual fielding of a successful and effective PMS system.

REFERENCES

Department of the Army, AR 602-1. (15 February 1983). Human factors engineering program. Washington, D.C.: Author.

Department of the Army, AR 602-2. (21 July 1986). Manpower and personnel integration (MANPRINT) in the materiel acquisition process. Washington, D.C.: Author.

Department of Defense, MIL-STD-1472C. (2 May 1981). Human engineering design criteria for military systems, equipment, and facilities. Washington, D.C.: Author.

McCracken, H., and Aldrich, T. (1984). Analyses of selected LHX mission functions (Army Research Institute Technical Note). Fort Rucker, AL: Army Research Institute Field Unit.

Stewart, J., and Shvern, U. (1987). HARDMAN-MIST analysis applied to the forward area air defense (FAAD) Pedestal-Mounted Stinger (PMS). Manuscript. Alexandria, VA: Army Research Institute, Manned Systems Group.

APPENDIX A

HUMAN FACTORS ENGINEERING (HFE)

(This section was prepared by the Army Materiel Test and Evaluation (ARMTE) Directorate, White Sands Missile Range, New Mexico. The material is as provided in the PMS NDICE Test Report, including section and page numbering. Principal author of this section was Mr. David Moran, ARMTE. ARI/HTI support personnel assisted as needed in the data collection phase.)

UNCLASSIFIED

(U) APPENDIX H (U)

HUMAN FACTORS ENGINEERING (HFE)

The attached was prepared as a Test Report input by the Army Materiel Test and Evaluation (ARMTE) Directorate, WSMR, NM, and corresponds to the HFE subtest in Appendix E of the PMS NDICE Test Design Plan.

UNCLASSIFIED

UNCLASSIFIED

THIS PAGE INTENTIONALLY LEFT BLANK

UNCLASSIFIED

H-2

UNCLASSIFIED

Human Factors Engineering

1 Objective

To evaluate the adequacy of the design of the PMS candidate systems to support operator functions.

2 Criteria

The fire unit will comply with applicable human engineering standards, AR 602-1, and AR 602-2 for design, performance, and operation. It is required that standards prescribed in MIL-STD-1472 be met. (ROC, App G, para 11(1)).

3 Data Acquisition Procedure

a. The Human Factors Engineering (HFE) evaluation was conducted in accordance with the guidelines of AR 602-1 and the TECOM Supplement thereto, and AR 602-2.

b. A HFE inspection was conducted on the PMS candidate systems. The arrangement, information content, and format of controls, displays, indicators, and labels were examined with regard to the requirements of MIL-STD-1472C and the HFE guidance of MIL-HDBK-759. All PMS controls were operated to assess whether excessive force or torque is required to activate any controls. Measurements were made of the illumination levels provided on the PMS gunner's control panels by the supplemental lighting when the system was in a dark environment.

c. An evaluation was made of whether the PMS candidate systems accommodate crew operations by soldiers between the 5th and 95th percentiles in body size for male ground troops. Soldiers representative of these body sizes performed gunner engagement tasks and driver tasks, as well as procedures for reloading missiles and reloading the predicted fire weapon, and procedures for converting Stinger missiles from the PMS configuration to the man-portable air defense system (MANPADS) configuration. Gunner engagement tasks, reloading of missiles and predicted fire weapons, and MANPADS conversion were also conducted by gunners wearing cold regions mittens. Operations by 5th and 95th percentile size operators and by operators wearing cold regions mittens were observed by HFE personnel and the operators were questioned concerning problems encountered.

d. In addition to the above described PMS activities which were conducted solely to gather HFE data, all of the other tracking, firing, and roadmarch activities of the Non-Developmental Item Candidate Evaluation (NDICE) test were

UNCLASSIFIED

UNCLASSIFIED

observed by HFE personnel. These activities included march orders; emplacements; tracking and firing from stationary PMS, moving PMS, and PMS configured for remote operation; missile and machine gun reloads; MANPADS conversions; and simulated emergency evacuation from the PMS candidates. These activities were performed during daytime and night by soldiers dressed either in battle dress uniforms or in MOPP IV outfits. Measurements were made of the times required to accomplish the various tasks and crewmen were interviewed concerning problems they encountered.

e. Steady state noise levels were measured at the gunner's and driver's ear locations with all noise producing equipment operating and with the vehicle both stationary and in motion in each of its forward gears.

f. Measurements were made of the horizontal and vertical visual angles available to the gunner from his operating position.

4 Results

a. The following are the results of the HFE inspections of the AVENGER, concerning conformance to MIL-STD-1472C and other HFE guidance.

(1) There are many differences regarding the arrangements of controls, switches, and indicators between the gunner's control console in the turret and the gunner's control console that is used for remote operations. Although the two control consoles cannot be identical, due to the smaller size of the remote console, it is desirable that the arrangement of the panels be made more consistent than they presently are. The lack of consistency between the control panels is a deviation from paragraph 4.2 and paragraph 5.1.2.1.1.4 of MIL-STD-1472C.

(2) On the gunner's console panels, all labels are above the associated controls, except the "focus" label, which is below the control. Paragraph 5.5.6.2.4 of MIL-STD-1472C states that labels are normally placed above controls. When a label is below a control, it can be obscured by the operators hand when he uses the control.

(3) The "level" label on the FLIR has partially worn off, so that some of the letters can no longer be read. All other labels have remained legible. Paragraph 5.5.4.5 of MIL-STD-1472C states that labels shall be mounted so as to minimize wear and shall remain legible.

(4) The FLIR polarity switch does not have a permanent label on the panel to indicate which switch position is black-hot

UNCLASSIFIED

UNCLASSIFIED

and which is white-hot. A temporary stick-on label with this information was placed on the panel at the crewmen's request when they first began operating the AVENGER. It is desirable that this information be labeled on the panel, even though it is obvious to any experienced AVENGER crewman whether his FLIR display is in black-hot or white-hot.

(5) On the FLIR control panel, the separation between the "gain" and "level" control knobs is only three quarters of an inch. Figure 7 of MIL-STD-1472C specifies a minimum separation distance of one inch for this type of knobs.

(6) The AVENGER legend lights are backlit only when activated, and the lettering is not legible when they are not activated. Paragraph 5.2.2.2.5 of MIL-STD-1472C states that lettering on single legend indicators shall be visible and legible whether or not the indicator is energized.

(7) The remote gunner's control console has a missile sequencing button, but unlike the gunner's console in the turret, the remote console does not display which of the eight missile positions is activated. This information might be important; for instance, if the AVENGER were carrying some basic Stingers and some Stinger POST missiles, the gunner might know which was which by their positions in the missile pods.

(8) On the remote gunner's control console the toggle switch position for "arm" is up and the position for "safe" is down. However, the labels on the split legend indicator are the reverse, with "arm" on the bottom and "safe" on the top. Paragraph 5.1.3.1 of MIL-STD-1472C states that display indicators shall clearly and unambiguously direct and guide the appropriate control response.

(9) The turret disable switch on the remote gunner's control console does not have an indicator to show that the turret is actually disabled, but shows only whether the toggle switch is in the "on" or "off" position. If the switch fails, there would be no indication that the turret was not disabled. Paragraph 5.2.2.1.2 of MIL-STD-1472C states that lights shall display equipment response, and not merely control position.

(10) The "laser enable" switch should have an associated indicator light to display whether the laser has been enabled or not. Paragraph 5.2.2.1.1 of MIL-STD-1472C states that transilluminated indicators should be used to draw attention to an important system status.

(11) The labeling on the circuit breaker panel would be improved if lines were placed on the panel to make it more clear

UNCLASSIFIED

UNCLASSIFIED

whether a label describes the circuit breaker above the label or the one below it. Paragraph 5.5.1.2 of MIL-STD-1472C states that label characteristics shall be consistent with the accuracy of identification required.

(12) There is no label on the FLIR control panel which says "FLIR Control Panel." Paragraph 5.5.6.2.4.d of MIL-STD-1472C states that functional groups shall be labeled.

(13) The FLIR power switch has positions for "on", "rain", and "off", and it is susceptible to being turned off inadvertently, when crewmen are attempting to switch between the other two positions. This has occurred at least three times during the NDICE test. Paragraph 5.4.1.8.1 of MIL-STD-1472C states that controls shall be designed so that they are not susceptible to being moved accidentally.

(14) The gunner's control console has many continuous adjustment rotary control knobs, including most of the FLIR control knobs. These knobs do not have pointers or markers to indicate the position the knob has been set to. It would be useful to have pointers or markers on these knobs, and also to have marks on the panel beside the knobs to indicate the full "on" and full "off" positions of the knobs. This would make it less likely that an operator would attempt to twist the knob beyond its stops, and also the operators could become acquainted with knob settings that provide good quality FLIR presentations for a variety of environmental conditions. Paragraph 5.4.2.2.1.1 of MIL-STD-1472C states that a pointer or marker should be provided on a continuous adjustment rotary control knob if control positions must be distinguished.

(15) The optical sight is located where it is susceptible to being kicked by a gunner entering or exiting the turret crew station, especially if the gunner is wearing NBC boots or cold regions boots. Paragraph 5.9.2.3 of MIL-STD-1472C states that delicate items shall be located or guarded so that they will not be susceptible to damage.

(16) A radio antenna is located right in front of the top of the ladder to the turret, so that it is an obstruction to entering the turret. The problem is alleviated somewhat by the fact that there is a standing platform to the right of the turret ladder, away from the antenna, and this platform is used by crewmen entering and exiting the turret. The antenna location presents an additional problem because it is often used as a handhold by gunners entering or exiting the turret. The antenna is inappropriate as a handhold because it is flexible and it is possible to receive an RF burn from the antenna if the radio is transmitting. Also, if the antenna is not installed on the

UNCLASSIFIED

UNCLASSIFIED

antenna post, the top of the antenna post can give an electrical shock to a person that touches it. Paragraph 5.13.4.2 of MIL-STD-1472C states that emergency doors and exits shall be unobstructed and shall not constitute a safety hazard.

(17) The gunner's seat in the turret is stood upon by the gunner as he enters and exits the gunner's station. If his boots are muddy the gunner's seat will be made muddy. Standing on the seat would also be expected to cause wear and tear on the seat, although the seat in the AVENGER which underwent NDICE testing has not exhibited excessive wear and tear. Paragraph 3.11.1 of MIL-HDBK-759A states that vehicle design should minimize the problems associated with introducing mud into crew areas.

(18) The crewmen have installed a makeshift rope lanyard on the turret canopy so that it is possible to easily pull the canopy closed. Such a lanyard should be installed on the vehicle as standard equipment. Paragraph 7.5.1.1.3 of MIL-HDBK-759A states that an operator should be able to easily latch a hatch door.

(19) The turret canopy has two latches, and requires two hand operation to operate, one hand for each latch. Paragraph 7.5.1.1.3 of MIL-HDBK-759A states that an operator should be able to latch a hatch in one continuous motion with one hand.

(20) If a gunner became incapacitated in the turret, it would be difficult to safely evacuate him from the turret, because the turret exit is so high above the ground. Paragraph 7.5.1.1 of MIL-HDBK-759A states that it is very important to assure that wounded or incapacitated crew members can be evacuated.

(21) The side to side width of the turret ladder rungs is 9 inches. Figure 36 of MIL-STD-1472C requires a minimum rung width of 12 inches. Climbing the ladder while wearing cold regions boots would be made more difficult by the inadequate rung width.

(22) The ventilator inlets in the turret have sharp edges, and crewmen can bump against them when entering or exiting the turret. Paragraph 5.13.5.4 of MIL-STD-1472C states that all exposed edges and corners shall be rounded.

(23) There are no armrests provided for the turret seat. Paragraph 5.7.3.4.5 of MIL-STD-1472C states that armrests shall be provided on chairs, unless otherwise specified.

(24) Some provisions of MIL-STD-1472C are not met in the AVENGER turret regarding seat height and console height, but the

UNCLASSIFIED

UNCLASSIFIED

AVENGER configuration is compatible with the activities that the gunner must perform. The gunner sits with his legs extended forward to reach his foot pedals, and because he does not sit with his lower leg straight up and down, he does not need for the seat to adjust upward to a height of 21 inches, as is required by paragraph 5.7.3.4.2 of MIL-STD-1472C. The AVENGER seat height adjusts from 12 inches to 15 inches above the turret floor, and this adjustment range is considered comfortable by gunners from 5th to 95th percentile in size. The shelf beneath the gunner's control console is only 22 inches above the floor, rather than the 25 inch kneeroom height specified by paragraph 5.7.3.5a of MIL-STD-1472C. However, because the seat only adjusts upward to a height of 15 inches, the console shelf height provides plenty of kneeroom for the operator if his legs are extended forward.

(25) The AVENGER driver's seat and the passenger seat in the cab are standard HMMWV seats and are government furnished equipment. These seats have no capability to be adjusted fore and aft, and therefore do not conform to paragraph 5.12.2.3 of MIL-STD-1472C, which requires 6 inches of fore and aft adjustment capability. Also, a 95th percentile size soldier reported that the driver's seat position caused his knees to be high and he thought he would become uncomfortable after extended periods there.

b. The following are the results of the HFE inspections of the CROSSBOW, concerning conformance to MIL-STD-1472C and other HFE guidance.

(1) Red lights are used on the gunner's control console to indicate missile status and to indicate activation or nonactivation of various subsystems such as FLIR, TV, IFF, laser rangefinder, etc. This is an improper use of the color red, according to paragraph 5.2.2.1.18 of MIL-STD-1472C. Red should be used for indicating no-go, failure, and hazardous and emergency situations. For the same reason, the boresight switch should not be red.

(2) Some of the labels of the gunner's controls are located below the associated control, such as "power-on and off," "stab enable," and "brk rlse." Paragraph 5.5.6.2.4b of MIL-STD-1472C states that labels are normally placed above the associated control. When a label is below a control it can be obscured by the operator's hand when he uses the control.

(3) The rotary control switch that releases the turret brake should have a directional arrow labeled beside it to show which direction to turn the switch to progressively release the brake. Paragraph 5.5.6.2.3c of MIL-STD-1472C states that control labeling shall indicate the functional result of control

UNCLASSIFIED

UNCLASSIFIED

movement. For the same reason, the crank which stows the turret for air transport, and unstows it, should also have a directional arrow labeled above the crank position to indicate which crank direction stows the turret.

(4) The turret brake release control extends over its label and partially obscures it. Paragraph 5.5.22 of MIL-STD-1472C states that controls should not obscure labels.

(5) The boresight control is not labeled, and the turret control thumbwheel on the gunner's gripstick is not labeled. Paragraph 5.5.1.1 of MIL-STD-1472C states that labels shall be provided whenever it is necessary for personnel to identify an item.

(6) On the turret control panel, the "on" and "off" positions are not labeled for the main power switch and the stabilization enable switch. Also, on the gunner's control console, there is no "on" and "off" position labeled for the toggle switches for "auto," "emer power off," and "frm avg." Also, these gunner's console switches are backwards, because the toggle switches are down for the "on" position and up for the "off" position. This is the reverse of the proper orientation, according to paragraph 5.4.1.2.1 of MIL-STD-1472C. Paragraph 5.5.6.2.3c indicates that the "on" and "off" positions should be labeled by the switches.

(7) The group of FLIR switches at the top of the gunner's control console should be labeled "FLIR". Paragraph 5.5.6.2.4d of MIL-STD-1472C states that functional groups shall be labeled.

(8) The label of the "arm" switch is located too far away from the "arm" switch so that it is not obvious that the label goes with the switch. Paragraph 5.5.2.2 of MIL-STD-1472C states that labels shall be placed on or very near the items which they identify.

(9) Many of the labels on the switches on the gunner's control console do not have adequate space between characters. The requirement, according to paragraph 5.5.5.11 of MIL-STD-1472C, is at least one stroke wide, but some of the letters on the CROSSBOW switches are right up against each other. Examples of switches with inadequate character spacing are: auto mode, manual uncage, sector cutout, auto track, hangfire, TV, lamp test, PTL set, and left set.

(10) The sector cutout and auto mode switch labels have line spacing of less than one-half character height, which violates paragraph 5.5.5.13 of MIL-STD-1472C.

UNCLASSIFIED

UNCLASSIFIED

(11) On the right side of the gunner's control console, the switch for enabling the laser range finder (LRF) is located directly above the IFF enable switch, but on the left side of the console, the associated switches are located in the reverse orientation, with IFF above LRF. Paragraph 5.1.2.1.1.4 of MIL-STD-1472C states that the location of recurring items shall be similar from panel to panel.

(12) The toggle switches at the top of the gunner's control console and the "pull man" rotary control at the bottom of the console are located where they are susceptible to damage, particularly when the console is removed and used remotely. A shield around the top of the panel would protect the toggles. Paragraph 5.9.2.3 of MIL-STD-1472C states that delicate items shall be located or guarded so that they will not be susceptible to damage while the unit is being handled.

(13) On the gunner's control console, the rotary control knobs labeled "pull man" and "pull bkup" should have labels which better describe their function. The "pull man" knob, when it is pulled out, provides a manual screen brightness adjustment, and when it is pushed in, it provides for automatic screen brightness adjustment. The "pull bkup" does not have a function to its pulled out position (it makes the screen go blank). In its pushed in position, it is a screen contrast control. Paragraph 5.5.3.1 of MIL-STD-1472C states that labels should primarily describe the functions of equipment items.

(14) The "pull man," "pull bkup," "sens," and "blk lvl" rotary controls should have pointers marked on the knobs and marks on the panel beside the knobs to indicate the full "on" and full "off" positions of the knobs. This would make it less likely that an operator would attempt to twist a knob beyond its stops. It would also allow operators to become acquainted with knob settings which provide good quality FLIR and TV presentations for a variety of environmental conditions. Paragraph 5.4.2.2.1.1 of MIL-STD-1472C states that a pointer or marker should be provided on a continuous adjustment rotary control knob if control positions must be distinguished.

(15) The handle on the top front of the gunner's control console does not have the required three inch depth for handles that will be used by mittened hands, as specified by Figure 48 of MIL-STD-1472C.

(16) The gunner's seat is at a height of 14 inches, and does not adjust in height. Paragraph 5.7.3.4.2 of MIL-STD-1472C states that work seating shall provide for seat height adjustment from 15 to 21 inches.

UNCLASSIFIED

UNCLASSIFIED

(17) The gunner's console is 18 inches above the floorboard of the CROSSBOW in the area above the gunner's knees. Paragraph 5.7.3.5a of MIL-STD-1472 specifies that kneeroom height for a seated work position shall be at least 25 inches. If the gunner extends his legs forward, as the CROSSBOW driver does, he can comfortably sit in the gunner's chair, but if he sits with his lower legs straight up and down he can bump the console with his knees. Moving the console higher would reduce the gunner's viewfield through the windshield.

(18) The CROSSBOW driver's seat is a standard HMMWV seat and is government furnished equipment. The seat has no capability to be adjusted fore and aft, and therefore does not conform to paragraph 5.12.2.3 of MIL-STD-1472C, which requires 6 inches of fore and aft adjustment capability. Also, a 95th percentile size soldier reported that the driver's seat position caused his knees to be high, and he thought he would become uncomfortable after extended periods there.

c. The only problem reported during AVENGER operations by representatives of the 5th and 95th percentile size soldiers is that the driver's seat position in the HMMWV cab is uncomfortable for 95th percentile size soldiers.

d. The only problem reported during CROSSBOW operations by representatives of the 5th and 95th percentile size soldiers is that the driver's seat position in the HMMWV cab is uncomfortable for 95th percentile size soldiers.

e. All AVENGER operations were demonstrated by an operator wearing cold regions mittens. All switch actions that would normally be performed using the index finger were performed using the thumb in the mittens. In order to operate the keypad of the termiflex, the operator had to use the eraser end of the pencil, grasped in the mitten, to press the keys.

f. All CROSSBOW operations were demonstrated by an operator wearing cold regions mittens. All switch actions that would normally be performed using the index finger were performed using the thumb in the mittens. The only problem reported was that the handle on the top of the gunner's control console does not accommodate a hand in a mitten.

g. The following are problems reported during interviews with AVENGER crewmen. In addition to crewman reports of test incidents, some crew comments are subjective, expressing opinions and preferences.

(1) During hot weather, the gunner's station in the turret can become very hot and, to a lesser extent, the driver's

UNCLASSIFIED

UNCLASSIFIED

station in the vehicle cab can also become uncomfortably hot. This discomfort is increased when the crewmen are wearing MOPP IV uniforms. There are no air conditioners on the AVENGER.

(2) On the remote gunner's control console, the IFF, auto slew, and FLIR buttons are not illuminated and their labels cannot be seen at night.

(3) The remote gunner's control console is heavy (approximately 65 pounds, according to contractor representatives) and is difficult to carry it the 50 meter distance from the vehicle to the remote gunner's location.

(4) The remote gunner's control console does not have the FLIR gain, focus, and level controls, and these controls must be adjusted in the turret crew station. The crewmen would like to have FLIR adjustment controls on the remote unit as well as in the turret.

(5) When operating the remote gunner's console, the gunner often has to assume uncomfortable body postures because he is sitting on the ground. The gunner's would like a stand for the remote that could adjust the height at which the console is positioned. The crewmen stated that when they moved the remote to be able to operate it more comfortably, they would then have to reintegrate the remote to the fire unit to account for the new orientation.

(6) The crewmen stated that wearing the MOPP IV mask with the laser eye shield sometimes made it difficult to see the screen symbology.

(7) The crewmen stated that reeling and unreeling the cable for the remote unit is difficult, and is a two man job.

(8) When there is dust on the turret canopy it becomes difficult to see targets through the optical sight. There is no means provided to clean the dust from the canopy. There is also no windshield wiper on the canopy for rain.

(9) The turret ventilator can draw dust into the gunner's crewstation especially if the filters are saturated.

(10) The canopy latches have occasionally popped open on bumpy roads.

(11) The antenna location at the top of the turret ladder makes it hard for crewmen to enter and exit the turret.

UNCLASSIFIED

UNCLASSIFIED

(12) The gunners would like to have the range and clock azimuth displayed on the FLIR screen rather than the termiflex so that they would not have to look away from the screen to obtain this information.

(13) The gunners would like to have the IFF interrogate button on the handstation, so that they could interrogate aircraft without removing a hand from the handstation.

(14) The crewmen report that sometimes either direct sunlight on the screen or glare from the sun's reflection off the canopy can wash out the FLIR presentation on the screen.

(15) The crewmen suggest that the pointer above the remote console be raised about a foot in height, so that the observer can operate the pointer without interfering with the gunner. Crewmen also suggest that the pointer be illuminated for night operations.

(16) The crewmen stated that their laser rangefinder sometimes went out of boresight alignment and did not provide target range returns.

(17) The machine gun magazine is heavy (approximately 68 pounds) and it is hard to align the magazine with its bracket.

(18) The gunners have reported that they have received minor cuts on their hands from both the FLIR receiver and the pod door when reloading the bottom missile. They have also cut their hands when reloading the machine gun.

h. The following are problems reported during interviews with CROSSBOW crewmen. In addition to crewman reports of test incidents, some crew comments are subjective, expressing opinions and preferences.

(1) It is possible to lose a target when switching field-of-view, because, in order to operate the field-of-view switch with his thumb he must momentarily remove his thumb from the thumb controller with which he tracks targets.

(2) The CROSSBOW radio antennas are located such that, as the pedestal rotates, the antenna can come between the FLIR receiver or TV camera and the target. When this occurs, the CROSSBOW can lose lockon of the target when the antenna blocks the target. It would also be possible to shoot a missile into the antenna and possibly damage the missile.

(3) The vibration of the vehicle on rough terrain during mobile engagements could make target acquisition difficult, auto

UNCLASSIFIED

UNCLASSIFIED

track impossible, manual tracking difficult, and laser range determination impossible.

(4) During hot weather the crewmen felt uncomfortably hot in the CROSSBOW cab, particularly when they were wearing NBC uniforms. There is no air conditioner in the CROSSBOW.

(5) The gunner must remove a cover from the laser manually. If he forgets to do so, he will not have the use of his laser.

(6) One of the gunners would prefer to have separate buttons for field-of-view change and for switching between FLIR and TV. Presently, both functions are on the same switch.

(7) When returning from remote operations of the gunner's console, retracting the remote cable is a difficult, slow, two man operation, with one man cranking the spool and one guiding the cable.

(8) The crewmen had difficulty in viewing the screen of the control console when they were wearing an NBC mask with laser eyeshield.

(9) The heaviness of the machine gun magazine makes it difficult to reload the machine gun.

1. The following are the times, expressed in seconds, for AVENGER crewmen standing outside the system to enter their crewstations, and the times for the crewmen to perform emergency exit from their crewstations:

(1) Time to enter both crew stations from standing on the ground in MOPP O.

7 7 6 6 7

median=7

(2) Time to emergency exit from both crew stations to the ground while dressed in MOPP O. During emergency exits from the gunner's turret, the gunner climbed from the turret onto the vehicle cab roof, and then jumped from the cab roof to the ground.

3 4 3 4 3

median=3

UNCLASSIFIED

UNCLASSIFIED

(3) Time to enter both crew stations from standing on the ground in MOPP IV.

10 9 11 10 5

median=10

(4) Time to emergency exit from both crew stations to the ground while dressed in MOPP IV.

10 7 4 6 6

median=6

j. The following are the times, expressed in seconds, for CROSSBOW crewmen standing outside the system to enter their crewstations, and the times for the crewmen to perform emergency exit from their crewstations.

(1) Time to enter both crew stations from standing outside the CROSSBOW dressed in MOPP O.

4 3 3 3 2

median=3

(2) Time to emergency exit from both crew stations while dressed in MOPP O.

2 2 2 3 2

median=2

(3) Time to enter both crew stations from standing outside the CROSSBOW dressed in MOPP IV.

15 16 5 5 4

median=5

(4) Time to emergency exit from both crew stations while dressed in MOPP IV.

5 4 4 7 3

median=4

k. Times are contained in previous subtests for emplacement and march order (para 3.1.8), missile and machine gun reloads

UNCLASSIFIED

UNCLASSIFIED

(para 3.1.9), and conversion to MANPADS configuration (para 3.3) for both AVENGER and CROSSBOW.

1. The following are the steady state noise levels measured at the AVENGER crewman's head positions. The sound levels are expressed in A-weighted decibels, which is abbreviated as dB(A).

(1) With vehicle stationary and all sound producing equipment on, the sound levels in the vehicle cab were:

	<u>driver's side</u>	<u>passenger side</u>
with windows up	81 dB(A)	81 dB(A)
with windows down	79 dB(A)	80 dB(A)

(2) With the vehicle stationary and all sound producing equipment on, the sound level at the gunner's station in the turret was 89 dB(A). With the heater ventilator off, and all other sound producing equipment on, the sound level was 62 dB(A).

(3) With the vehicle moving and all sound producing equipment on, and the windows rolled up, the sound levels at the driver's station were:

First gear, gravel road	84 dB(A)
Second gear, gravel road	84 dB(A)
Drive gear, gravel road	85 dB(A)
First gear, paved road	84 dB(A)
Second gear, paved road	87 dB(A)
Drive gear, paved road	89 dB(A)

(4) With the vehicle moving and all sound producing equipment on, including the gunner's heater ventilator, the sound levels at the gunner's station were:

First gear, gravel road	89 dB(A)
Second gear, gravel road	89 dB(A)
Drive gear, gravel road	89 dB(A)
First gear, paved road	90 dB(A)
Second gear, paved road	89 dB(A)

UNCLASSIFIED

UNCLASSIFIED

Drive gear, paved road 90 dB(A)

(5) With the gunner's heater ventilator off and all other sound producing equipment on, the sound levels in the turret, while moving over paved and gravel roads, did not exceed 80 dB(A). Most of the noise in the gunner's station is produced by the heater ventilator.

1. The following are the steady state noise levels measured at the CROSSBOW gunner's head position. The sound levels at the driver's position were virtually the same.

(1) With the vehicle stationary and all sound producing equipment on, the sound level was 75 dB(A), both with the windows up and with the windows down.

(2) With the vehicle moving, and all sound producing equipment on, and the windows rolled up, the sound levels were:

First, gear, gravel road	82 dB(A)
Second gear, gravel road	86 dB(A)
Drive gear, gravel road	85 dB(A)
First gear, paved road	83 dB(A)
Second gear, paved road	84 dB(A)
Drive gear, paved road	87 dB(A)

m. The following are the approximate viewing angles of the AVENGER gunner from the turret, measured from the eye level of a 50th percentile size soldier.

(1) The horizontal viewing angle extends 43 degrees to each side of the turret centerline, for a total horizontal viewing angle of 86 degrees. The position of the missile pods blocks off further peripheral viewing by the gunner.

(2) The downward vertical viewing angle extends 10 degrees downward from the gunner eye level.

(3) The upward vertical viewing angle extends 90 degrees above the gunner's eye level. The total vertical viewing angle is 100 degrees.

n. The following are the approximate viewing angles of CROSSBOW gunner from his crew station in the vehicle, measured from the eye level of a 50 percentile size soldier.

UNCLASSIFIED

UNCLASSIFIED

(1) The horizontal viewing angle through the front windshield extends from 25 degrees to the right of straight ahead to 60 degrees to the left of straight ahead, with 15 degrees of this viewfield blocked off by the centerpost of the windshield. The horizontal viewing angle out the side window extends from 45 degrees to the right of straight ahead to 90 degrees to the right of straight ahead. The horizontal viewfield through the windshield encompasses 70 degrees and the horizontal viewfield out the side window encompasses 45 degrees.

(2) The vertical viewfield through the windshield extends from 5 degrees below horizontal to 25 degrees above horizontal. The vertical viewfield through the side window extends from 15 degrees below horizontal to 15 degrees above horizontal.

p. The illumination levels measured for the AVENGER auxiliary white light, under dark background conditions, was 34 footcandles one inch in front of the light and one to five footcandles at the operating gunner's console.

q. The illumination levels measured for the CROSSBOW auxiliary blue light, under dark background conditions, was 58 footcandles one inch away from the light and two to five footcandles at the operating gunner's console. The small mobile spotlight of the CROSSBOW was not operational during the lighting test.

r. The workspace lighting levels described in MIL-STD-1472C do not apply because FLIR operations are best performed under low level lighting conditions.

5 Analysis

a. The criteria were met with regard to the HFE evaluation being conducted according to the guidelines of AR 602-1 and AR 602-2.

b. The criteria were partially met with regard to the PMS systems conforming to the provisions of MIL-STD-1472. The areas in which the systems did not conform to MIL-STD-1472 are described in paragraphs 4a and 4b above.

c. When the system is stationary, the steady state noise levels in the AVENGER vehicle cab are low enough that no hearing protection is required, according to MIL-STD-1474. However, in the gunner's station in the turret, the noise levels in the stationary mode require single hearing protection whenever the heater ventilator is operating. When the system is mobile, the steady state noise levels are high enough that single hearing

UNCLASSIFIED

UNCLASSIFIED

protection is required in the vehicle cab, but not in the gunner's station if the heater ventilator is not operating. Adequate single hearing protection is provided the DH-132 communications helmet used on AVENGER.

d. When the system is stationary, the steady state noise levels of the CROSSBOW are low enough that no hearing protection is required, according to MIL-STD-1474. When the system is mobile, the steady state noise levels are high enough that single hearing protection is required. The DH-132 communications helmet used on the system provides adequate attenuation of these sound levels to produce a safe hearing environment.

UNCLASSIFIED

APPENDIX B

SYSTEM SAFETY AND HEALTH HAZARDS

(This section was prepared by the Army Materiel Test and Evaluation (ARMTE) Directorate, White Sands Missile Range, New Mexico. The material is as provided in the PMS NDICE Test Report, including section and page numbering. Principal author was Mr. David Moran, ARMTE. ARI/HTI support personnel assisted as needed in the data collection phase.)

UNCLASSIFIED

(U) APPENDIX E (U)

SAFETY AND HEALTH HAZARDS

The attached was prepared as a Test Report input by the Army Materiel Test and Evaluation (ARMTE) Directorate, WSMR, NM, and corresponds to the Safety subtest in Appendix E of the PMS NDICE Test Design Plan.

UNCLASSIFIED

UNCLASSIFIED

THIS PAGE INTENTIONALLY LEFT BLANK

UNCLASSIFIED

UNCLASSIFIED

SAFETY AND HEALTH HAZARDS

1 Objective

To identify any hazards associated with the operations of the PMS systems.

2 Criteria

a. The following criteria are derived from MIS-19804 and are listed along with their paragraph numbers in that document. A partial evaluation of most of these criteria was accomplished during this safety test.

b. The system shall suffer no physical damage and shall meet the performance requirements after exposure to the pressure loads of the Stinger missile plume. (MIS-19804, para 3.2.7.2.2.4)

c. The PMS system shall be compatible with and shall not degrade the safety features of the Government-furnished property (GFP) elements involved. Safety features shall provide for maximum safety and protection of operating and maintenance personnel and associated equipment. Design and safety verification shall be accomplished in accordance with the safety criteria contained in MIL-STD-882 and conform to the human engineering design criteria of MIL-STD-1472. The system shall meet the following safety requirements. (MIS-19804, para 3.3.5)

d. The system shall meet MIL-STD-454, requirement 1, with regard to electrical/electronic safety and requirement 27 with regard to battery compartment design. (MIS-19804, para 3.3.5.2)

e. Personnel shall not be exposed to hazardous materials in excess of the National Institute for Occupational Safety and Health/Occupational Safety and Health Act (NIOSH/OSHA) established permissible exposure limits during normal operations. (MIS-19804, para 3.3.5.3)

f. Personnel shall be protected from inadvertent contact with moving mechanical parts, sharp corners, edges, and projections with hazardous characteristics during operation, test or maintenance. Guards or other safety devices shall be provided. (MIS-19804, para 3.3.5.4a)

g. A seat restraining system shall be provided for the gunner. (MIS-19804, para 3.3.5.4d)

UNCLASSIFIED

UNCLASSIFIED

h. Control switches shall be designed, located, and positioned to minimize the probability of inadvertent activation. (MIS-19804, para 3.3.5.4e)

i. The system design shall ensure that it is mechanically or electrically impossible to activate controls in improper sequence or connect components and subsystems improperly. (MIS-19804, para 3.3.5.4f)

j. Multiple sequential actions shall be required to launch the missile. Weapon operating controls and circuits shall be designed to prevent firing when improper operation sequence or improper control operation is undertaken. (MIS-19804, para 3.3.5.4g)

k. Power and stored energy sources shall be isolated from fire controls and circuits until intentionally activated. (MIS-19804, para 3.3.5.4h)

l. The PMS system access doors or hatches shall conform to MIL-STD-1472C which are readily accessible, unobstructed and provided with quick-opening devices that operate in 3 seconds or less. (MIS-19804, para 3.3.5.5)

m. The system shall not adversely subject the operating personnel/vehicle to blast, noise, heat, debris, or toxicity from Stinger missile launch and flight motor firings with the application of Stinger weapon rounds in the PMS system configuration. The noise levels for PMS system and related equipment shall not exceed that specified in MIL-STD-1474. (MIS-19804, para 3.3.5.8)

n. Safety of laser range-finder devices shall conform to the provisions of MIL-STD-454 and TB-MED-279. (Test Agency Criterion)

3 Data Acquisition Procedure

a. The Safety Assessment Reports of the AVENGER and the CROSSBOW were reviewed by a system safety engineer (SSE) and a human factors engineer (HFE). A safety inspection of the AVENGER and the CROSSBOW was conducted by the SSE and checklists were completed concerning compliance with MIL-STD-454, Standard General Requirements for Electronic Equipment, and TOP 10-2-508, Safety and Health Evaluation - General Equipment.

b. Measurements were made of toxic gas levels entering the crew stations from the vehicle diesel exhaust and, in the case of AVENGER, from the diesel exhaust of the turret heater/ventilator. Measurements were made of levels of carbon monoxide (CO),

UNCLASSIFIED

UNCLASSIFIED

nitrogen dioxide (NO₂) and formaldehyde. The measurements were conducted under various combinations of conditions, involving vehicle doors and windows open or closed, EMMV ventilator system on or off, and, for AVENGER, with the turret heater/ventilator on or off. Under each condition, toxic gas measurements were conducted for a period of 20 to 30 minutes. The ambient wind did not exceed 10 miles per hour during the test.

c. In order to gather the data needed to evaluate crewman safety during the firing of Stinger missiles, toxic gas and noise measurements were obtained at the crewstations during live firings conducted by contractor personnel remotely operating from a distance of 50 meters from the vehicle.

(1) The missiles fired during the safety tests were not actual Stinger missiles, but rather they were Ballistic Test Vehicles (BTVs). A BTV provides a simulation of the blast and exhaust effects of a Stinger missile because a BTV contains the same launch motor as a Stinger missile, although it does not have a flight motor as a Stinger missile does.

(2) For both the AVENGER and the CROSSBOW, the BTV firings were conducted with the vehicle in a stationary position with the system power on, the doors and windows closed, the vehicle cab ventilator off, and the vehicle cab ventilator inlets closed. Additionally, in the case of the AVENGER, the gunner's station ventilator was on. The gunner's station ventilator in the turret automatically closes its inlets immediately before a missile is fired. The AVENGER vehicle engine was on during the test, and the CROSSBOW vehicle engine was off. The gunner's fire controls for both the AVENGER and the CROSSBOW were operated from a remote firing location, and the opening through which the gunner control cables extended from the vehicle was sealed with tape.

(3) Except for one firing by AVENGER and one by CROSSBOW, the launch angle of each BTV was oriented so as to direct a worst-case blast impingement on the upper rear portion of the right side vehicle cab door. One BTV for each system was fired at a best-case orientation, in that its azimuth angle was perpendicular to the direction of the vehicle and there was no blast impingement on the vehicle cab.

(4) The toxic gases which were measured at the crewstations during BTV firings were hydrogen chloride (HCl) and carbon monoxide (CO). The HCl gas was measured by means of a sequential impinger device and the CO was measured by means of an Interscan CO meter with chart recorder.

UNCLASSIFIED

UNCLASSIFIED

(5) The impulse noise levels from BTV firings were measured at all crewstations by means of microphones at head level. Impulse noise levels were also measured external to the vehicle to determine the 140 decibel (dB) contour around the vehicle, which is the area within which hearing protection must be worn during firings.

d. Toxic gas and noise measurements were also obtained during live firings of the machine guns of the AVENGER and CROSSBOW. The firings were conducted by contractor personnel remotely operating from a distance of 50 meters from the vehicle.

(1) The conditions of the AVENGER during the test were with the vehicle in the stationary position with the vehicle engine on, the windows closed, and the ventilators in both the vehicle cab and the turret turned on with the ventilator inlets open.

(2) The conditions of the CROSSBOW during the test were with the vehicle in the stationary position with the vehicle engine off, the system power on, the windows and doors closed, the ventilator off, and the ventilator inlets closed.

(3) Each time the test was conducted, 200 rounds of ammunition were fired, with the gun level and pointed over the driver's door.

(4) The toxic gases measured during the machine gun firings were carbon monoxide (CO), nitrogen dioxide (NO₂) and sulfur dioxide (SO₂). The measurements were made at all crewstations with Interscan monitors with chart recorders.

(5) The impulse noise measurements during the machine gun firings were conducted in the same manner and with the same microphone locations as the measurements during the BTV firings, as described above.

e. Wind speeds were measured during the BTV and machine gun firings.

f. During environmental chamber testing at a dry bulb ambient temperature of approximately 140 degrees F (60 degrees C), the effective temperature was derived from the measured dry bulb and wet bulb temperatures and the ventilator airspeed at the AVENGER and CROSSBOW crewman locations. The effective temperatures were derived in accordance with Figure 39 of MIL-STD-1472C.

g. Descriptions of safety related incidents and observations were obtained from Test Incident Reports (TIRs), human factors

UNCLASSIFIED

UNCLASSIFIED

inspections regarding compliance to MIL-STD-1472C, and crewman comments regarding safety.

4 Results

a. The results of the review of the Safety Assessment Report and the safety inspections of the systems are as follows:

(1) The AVENGER complies with the safety requirements of MIL-STD-454 and TOP 10-2-508 with the exception of the AVENGER safety problems described below. (Paragraph aa)

(2) The CROSSBOW complies with the safety requirements of MIL-STD-454 and TOP 1-2-508, with the exception of the CROSSBOW safety problems described below. (Paragraph aa)

b. During the test for toxic gas from vehicle exhaust, the CO in the AVENGER crewstations did not exceed 2 parts per million (ppm), the NO₂ level did not exceed 0.2 ppm, and no formaldehyde was detected. The OSHA limit for long term exposure to CO is 50 ppm, and the limit for NO₂ is 5 ppm.

c. During the test for toxic gases from vehicle exhaust, the CO level in the CROSSBOW crewstations did not exceed 6 ppm, no NO₂ was detected, and no formaldehyde was detected.

d. The following are the HCl levels measured in the AVENGER vehicle cab for the BTV firings oriented to produce worst case blast impingement on the vehicle cab doors (all BTVs except BTV #6). The HCl levels are expressed in milligrams per cubic meter (mg/m³).

(1) BTV #1		HCl Concentration
<u>Sample #</u>	<u>Time Interval in seconds</u>	<u>in mg/m³</u>
1	from(t=0) to (t=+5)	346
2	from(t=+5) to (t=+20)	144
3	from(t=+20) to (t=+35)	206
4	from(t=+35) to (t=+50)	82
5	from(t=+50) to (t=+65)	74
6	from(t=+65) to (t=+80)	45
7	from(t=+80) to (t=+95)	58

UNCLASSIFIED

UNCLASSIFIED

(2) BTV #2 & #3 (two round ripple firing)

<u>Sample #</u>	<u>Time Interval in seconds</u>	<u>HCl Concentration in mg/m³</u>
1	from(t=0) to (t=+5)	1296
2	from(t=+5) to (t=+20)	54
3	from(t=+20) to (t=+35)	29
4	from(t=+35) to (t=+50)	16
5	from(t=+50) to (t=+65)	14
6	from(t=+65) to (t=+80)	7
7	from(t=+80) to (t=+95)	41

(3) BTV #4 & #5 (two round ripple firing)

<u>Sample #</u>	<u>Time Interval in seconds</u>	<u>HCl Concentration in mg/m³</u>
1	from(t=0) to (t=+5)	309
2	from(t=+5) to (t=+20)	16
3-7	from(t=+20) to (t=+95)	none detected

(4) BTV #7 & #8 (two round ripple firing)

<u>Sample #</u>	<u>Time Interval in seconds</u>	<u>HCl Concentration in mg/m³</u>
1	from(t=0) to (t=+15)	16
2	from(t=+15) to (t=+30)	42
3	from(t=+30) to (t=+45)	90
4	from(t=+45) to (t=+60)	160
5	from(t=+60) to (t=+75)	239
6	from(t=+75) to (t=+90)	185
7	from(t=+90) to (t=+105)	173
8	from(t=+105) to (t=+120)	119

(5) During the two round ripple firing of BTV #2 and #3 and the two round ripple firing of BTV #4 and #5, the missile backblast forced open the doors of the AVENGER Cab, and caused damage to the cab doors. After these firings, and before the firing of BTVs #6, #7, and #8, blast deflectors and door seal modifications were installed to the AVENGER to prevent the the doors from being forced open.

e. The HCl levels at the AVENGER gunner's station in the turret were measured only during the firing of BTVs 1 through 5. No HCl was detected in the gunner's station in the turret during the test.

UNCLASSIFIED

UNCLASSIFIED

f. Measurements of HCl were conducted in the AVENGER vehicle cab during the firing of BTV #6, which was a best-case rather than a worst-case condition, because the firing angle did not impinge the backblast on the vehicle cab. Measured clockwise from the direction of the vehicle, BTV #6 had a firing azimuth of 270 degrees and a firing elevation of 14 degrees. No HCl was detected for BTV #6.

g. The following are the CO levels, expressed in ppm, measured in the AVENGER cab during the BTV firings. Due to an instrumentation malfunction, no CO data were obtained at the gunner's station.

<u>BTV #</u>	<u>CO level</u>
1	none detected
2&3 (two round ripple)	5 ppm
4&5	none detected
6	15 ppm
7&8 (two round ripple)	200 ppm

h. The impulse noise levels from the BTV firings were measured in the AVENGER turret and the vehicle cab by means of microphones at head level at the gunner and driver and passenger stations. Impulse noise levels were also measured external to the vehicle to determine the 140 decibel (dB) contour around the vehicle, which is the area within which hearing protection must be worn during firings. The following are the impulse noise levels, in the decibels, in the vehicle cab. Also listed is the B-Duration of the impulse noise, which is the length of time in milliseconds (ms) that the sound level is within 20 dB of the peak.

	<u>Impulse Noise</u> <u>in dB</u>	<u>B-Duration</u> <u>in ms</u>
(1) BTV #1		
driver's ear	158.6	96
passenger's ear	157.8	101
gunner's ear	144.9	105
(2) BTV #2		
driver's ear	163.8	93
passenger's ear	163.5	95
gunner's ear	144.0	114

UNCLASSIFIED

UNCLASSIFIED

(3) BTV #3

driver's ear	above 164	84
	(microphone saturated)	
passenger's ear	above 164	82
gunner's ear	149.8	104

(4) BTV #6

driver's ear	146.5	36
passenger's ear	146.5	35

(5) BTV #7

driver's ear	163.9	90
passenger's ear	163.9	87

(6) BTV #8

driver's ear	163.0	92
passenger's ear	166.1	95

i. The maximum noise measured external to the AVENGER vehicle was directly behind the BTV backblast during BTV #1, which has a 140 dB contour extending 575 feet from the vehicle. Because the AVENGER can fire in any direction, the 140 dB contour around the AVENGER is a circle with a radius of 575 feet.

j. The following are the measured wind speeds in miles per hour (mph) measured during the AVENGER BTV firings:

<u>BTV#</u>	<u>wind speed (mph)</u>
1	13
2&3 (two round ripple)	18
4&5 (two round ripple)	14
6	11
7&8 (two round ripple)	10

k. The following are the HCl levels measured in the CROSSBOW for the BTV firings oriented to produce worst case blast impingement on the vehicle cab doors (all BTVs except BTV #2). The HCl levels are expressed in milligrams per cubic meter (mg/m^3). BTV #3 and #4 were slightly misdirected, so they failed to fully achieve the intended worst-case blast impingement on the cab door.

UNCLASSIFIED

UNCLASSIFIED

(1) BTV #1

<u>Sample #</u>	<u>Time Interval in seconds</u>	<u>HCl Concentration in mg/m^3</u>
1	from(t=0) to (t=+5)	21
2	from(t=+5) to (t=+20)	17
3	from(t=+20) to (t=+35)	16
4	from(t=+35) to (t=+50)	144
5	from(t=+50) to (t=+65)	148
6	from(t=+65) to (t=+80)	105
7	from(t=+80) to (t=+95)	14

(2) BTV #3 & #4 (two round ripple firing)

<u>Sample #</u>	<u>Time Interval in seconds</u>	<u>HCl Concentration in mg/m^3</u>
1	from(t=0) to (t=+15)	None Detected
2	from(t=+15) to (t=+30)	13
3	from(t=+30) to (t=+45)	14
4	from(t=+45) to (t=+60)	13
5	from(t=+60) to (t=+75)	13
6	from(t=+75) to (t=+90)	11
7	from(t=+90) to (t=+105)	11
8	from(t=+105) to (t=+120)	11

(3) BTV #5 & #6 (two round ripple firing)

<u>Gunner's Seat Sample #</u>	<u>Time Interval in seconds</u>	<u>HCl Concentration in mg/m^3</u>
1	from(t=0) to (t=+15)	66
2	from(t=+15) to (t=+30)	78
3	from(t=+30) to (t=+45)	103
4	from(t=+45) to (t=+60)	26
5	from(t=+60) to (t=+75)	17
6	from(t=+75) to (t=+90)	10
7	from(t=+90) to (t=+105)	7
8	from(t=+105) to (t=+120)	7

<u>Driver's Seat Sample #</u>	<u>Time Interval in seconds</u>	<u>HCl Concentration in mg/m^3</u>
1	from(t=0) to (t=+15)	132
2	from(t=+15) to (t=+30)	177
3	from(t=+30) to (t=+45)	79

UNCLASSIFIED

UNCLASSIFIED

4	from(t=+45) to (t=+60)	43
5	from(t=+60) to (t=+75)	23
6	from(t=+75) to (t=+90)	21
7	from(t=+90) to (t=+105)	17
8	from(t=+105) to (t=+120)	14

(4) BTV #7 & #8 (two round ripple firing)

<u>Sample #</u>	<u>Time Interval in seconds</u>	<u>HCl Concentration in mg/m³</u>
1	from(t=0) to (t=+10)	22
2	from(t=+10) to (t=+25)	10
3	from(t=+25) to (t=+40)	10
4	from(t=+40) to (t=+55)	7
5	from(t=+45) to (t=+60)	3
6	from(t=+60) to (t=+75)	3
7	from(t=+75) to (t=+90)	3
8	from(t=+90) to (t=+105)	None Detected

(5) BTV #9 & #10 (two round ripple firing)

<u>Sample #</u>	<u>Time Interval in seconds</u>	<u>HCl Concentration in mg/m³</u>
1	from(t=0) to (t=+5)	22
2	from(t=+5) to (t=+20)	14
3	from(t=+20) to (t=+35)	10
4	from(t=+35) to (t=+50)	10
5	from(t=+50) to (t=+65)	7
6	from(t=+65) to (t=+80)	7
7	from(t=+80) to (t=+95)	3
8	from(t=+95) to (t=+110)	3

1. No HCl gas was detected for BTV #2, which had a firing angle which did not impinge on the vehicle cab. Measured clockwise from the direction of the vehicle, BTV #2 had a firing azimuth of 270 degrees and a firing elevation of 14 degrees.

m. The following are the CO levels, expressed in parts per million (ppm), measured during the BTV firings from the CROSSBOW

<u>BTV #</u>	<u>CO level (ppm)</u>
1	1
2	25
3&4	15
5&6	440 level returned to baseline in 3 minutes
7&8	13
9&10	25

UNCLASSIFIED

UNCLASSIFIED

n. During the firing of CROSSBOW BTV #1 and the ripple firing of BTV #5 and BTV #6, the missile backblast forced open the gunner's side cab door, and caused extensive damage to the door. During BTV #1 the driver's side door was also forced open and damaged by the force of the blast. During the period after the firing of BTV #1 and before the other BTV firings, blast deflectors were installed to prevent the missile backblast from blowing the door open. The results of the ripple firing of BTV #5 and #6 indicate that these modifications were not successful. During the period after the firing of BTV #6 and before BTV #7, the blast deflectors were modified and additional latches were placed on the cab doors. Two types of doors were used during the last four BTV firings, a rigid door for BTV #7 and #8, and the usual CROSSBOW flexible door for BTV #9 and #10. Both types of doors stayed closed and intact during these firings, and HCl level were decreased from levels during previous BTV firings at worst-case launch angles.

o. The impulse noise levels from the CROSSBOW BTV firings were measured in the vehicle cab by means of microphones at head level at the gunner and driver stations. Impulse noise levels were also measured external to the vehicle to determine the 140 decibel (dB) contour around the vehicle, which is the area within which hearing protection must be worn during firings. The following are the impulse noise levels, in decibels, in the vehicle cab. Also listed is the B-Duration of the impulse noise, which is the length of time in milliseconds (ms) that the sound level is within 20 dB of the peak. Impulse noise levels were not measured during BTV #5 and #6.

	<u>Impulse Noise</u> <u>in dB</u>	<u>B-Duration</u> <u>in ms</u>
(1) BTV #1		
driver's ear	170.1	65
gunner's ear	172.1	69
(2) BTV #2		
driver's ear	145.6	64
gunner's ear	147.0	58
(3) BTV #3		
driver's ear	157.6	75
gunner's ear	156.9	70

UNCLASSIFIED

UNCLASSIFIED

(4) BTV #4

driver's ear	158.3	82
gunner's ear	160.3	84

p. The maximum noise measured external to the CROSSBOW vehicle was directly behind the BTV backblast during BTV #1, which had a 140 dB contour extending 519 feet from the vehicle. Because the CROSSBOW can fire in any direction, the 140 dB contour around the CROSSBOW is a circle with a radius of 519 feet.

q. The following are the measured wind speeds during the CROSSBOW BTV firings:

<u>BTV#</u>	<u>wind speed(mph)</u>
1	2
2	2
3 & 4	6
5 & 6	4
7 & 8	8
9 & 10	9

r. The following are the results of the toxic gas tests of the AVENGER machine gun, which is also known as the predicted fire weapon (PFW).

(1) First test at gunner's station had no valid data due to an instrumentation malfunction.

(2) Second test at gunner's station	Concentration level
CO	none detected
SO2	none detected
NO2	none detected

(3) Test at driver's station	
CO	3 ppm
SO2	3 ppm
NO2	0.3 ppm

s. The following are the impulse noise measurements from AVENGER PFW firing measured at the crewmen's head location:

driver's ear	no data (microphone malfunction)
passenger's ear	146.3 dB
gunner's ear	149.8 dB

UNCLASSIFIED

UNCLASSIFIED

t. The maximum noise measured external to the AVENGER vehicle was along the line of fire, which had a 140 dB contour extending 288 feet from the vehicle. Because the AVENGER can fire in any direction, the 140 dB contour around the AVENGER is a circle with a radius of 288 feet.

u. The following are the measured winds speeds during the AVENGER PFW firings:

(1) During the cab toxic gas test, the first turret toxic gas test, and the noise test, the wind varied from 4 to 8 mph.

(2) During the second toxic gas test in the turret (gunner's station), the wind was 4 mph.

v. The following are the results of the toxic gas tests during CROSSBOW PFW firing, expressed in ppm:

	Concentration level
(a) First PFW test	
CO	1 ppm
NO ₂	none detected
SO ₂	instrumentation malfunction
(b) Second PFW test	
CO	above 100 ppm*for 20 seconds
NO ₂	instrumentation malfunction
SO ₂	29 ppm
(c) Third PFW test	
CO	instrumentation malfunction
NO ₂	2 ppm
SO ₂	1 ppm

*The CO level went off the meter scale at 100 ppm.

w. The following are the impulse noise measurements from PFW firing measured at the crewmen's head locations:

driver's ear	151.6 dB
gunner's ear	148.6 dB

x. The maximum PFW noise measured external to the vehicle was along the line of fire, which had a 140 dB contour extending 288 feet from the vehicle. Because the CROSSBOW can fire in any

UNCLASSIFIED

UNCLASSIFIED

direction, the 140 dB contour around the CROSSBOW is a circle with a radius of 288 feet.

y. The following are the measured wind speeds during the PFW firings:

- (1) First PFW test 19 to 30 mph
- (2) Second PFW test 5 mph
- (3) Third PFW test 4 mph

z. The following are the results of the test of effective temperature in the hot environmental chamber, with an ambient dry bulb temperature which varied between 140 degrees F (60 degrees C) and 145 degrees F (63 degrees C).

(1) The measured ventilator airspeeds are as follows:

(a) At the crewman's head position in the AVENGER turret - 50 feet per minute.

(b) At the crewman's head position in the AVENGER vehicle cab - 0 to 15 feet per minute.

(c) At the crewman's head position in the CROSSBOW vehicle cab - 0 to 15 feet per minute.

(2) The chamber wet bulb temperature varied from 86 degrees F (31 degrees C) to 88 degrees F (31 degrees C).

(3) The chamber humidity was 12 percent.

(4) Because the crewstations are not air conditioned, the effective temperature is essentially determined by the chamber dry bulb temperature and wet bulb temperature. According to the nomograph presented on Figure 39 of MIL-STD-1472, the effective temperature at all crewstations in the chamber varied from 99 degrees F (37 degrees C) to 102 degrees F (39 degrees C). For crewmen dressed in MOPP IV, effective temperatures increase by 10 degrees F (5 degrees C), according to paragraph 5.8.1.8 of MIL-STD-1472C. Therefore the effective temperatures at all crewstations for crewmen in MOPP IV attire varied from 109 degrees F (42 degrees C) to 112 degrees F (44 degrees C) under the conditions in the hot environmental chamber.

aa. The following are the safety problems which were reported through human factors engineering (HFE) inspections, crewman interviews, and TIRs. For those problems which are also

UNCLASSIFIED

UNCLASSIFIED

mentioned in the HFE subtest report, the paragraph number from that report is noted with the problem descriptions below.

(1) The radio antenna at the entryway to the AVENGER turret is often used as a handhold by gunners entering and exiting the turret. The antenna is too flexible to be a suitable handle and, if the radio is transmitting, the antenna can give an RF burn to a crewman's hand when he grabs it. The antenna post, when the antenna is removed, can give an electrical shock to a crewman's hand if he grabs the top of the post. (Refer to Appendix H para 4a(16)).

(2) The AVENGER gunners have reported receiving minor cuts on their hands from both the FLIR receiver and pod door when reloading the bottom missile. They have also cut their hands during reload of the machine gun. (Para 2.6.4.g(18), Appendix H reports the same problem).

(3) The ventilator inlets in the AVENGER turret have sharp edges, and crewmen can bump against them when entering or exiting the turret. (The same problem is reported in Appendix H para 4a (22)).

(4) The CROSSBOW system is capable of firing a missile into the radio antennas, and possibly damaging the missile. (This problem is also reported in Appendix H para 4h(2)).

(5) During replacement of HMMWV batteries, as the battery cable is being connected to the terminals, it is possible to cause some nearby filter capacitors to discharge and spark. Because these capacitors hold a charge of only 2 volts each, they do not present a significant personnel hazard, but the sparking could cause damage to the HMMWV battery.

5 Analysis

a. The following analysis applies to the toxic gas and noise levels measured in the AVENGER during BTV firings.

(1) Hydrogen chloride concentration limits are specified by two federal government agencies, the Occupational Safety and Health Administration (OSHA) and the National Institute for Occupational Safety and Health (NIOSH). The HCl limit specified by OSHA is 7 mg/m³ or 5 ppm, and that limit is intended to apply regardless of the duration of exposure. However, NIOSH specifies a HCl limit of 140 mg/m³ or 100 ppm for short duration exposures of 30 minutes or less. Both the OSHA and NIOSH limits for HCl were exceeded in the cab during every BTV firing except BTV #6, which was also the only BTV which did not have the missile

UNCLASSIFIED

UNCLASSIFIED

backblast directed at the cab. No HCl gas was detected during the firing of BTV #6.

(2) The blast deflectors and door sealing modifications which were installed on the AVENGER before the firing of BTVs #6, #7, and #8 succeeded in preventing the cab doors from being forced open during firings.

(3) No HCl was detected in the gunner's station in the turret.

(4) The 200 ppm level of CO measured during the two round ripple firing of BTVs #7 and #8 is equal to the NIOSH ceiling limit, or maximum allowable concentration, of 200 ppm CO.

(5) According to Figure 5 of MIL-STD-1474, the impulse noise levels and B-Durations measured in the vehicle cab during BTV firings are allowable if the crew wears single hearing protection. Sound attenuating communications helmets are considered an adequate form of single hearing protection.

(6) The 140 dB contour around the AVENGER, which is the area in which hearing protection must be worn, is determined by the BTV noise rather than by the machine gun noise, because a BTV is louder than the machine gun. The 140 dB contour is a circle around the AVENGER with a radius of 575 feet. In a situation in which the machine gun will be fired but no missile will be fired, the 140 dB contour is a circle around the AVENGER with a radius of 288 feet.

b. The following analysis applies to the toxic gas and noise levels measured in the CROSSBOW during BTV firings.

(1) The OSHA limit for HCl, which is 7 mg/m³, was exceeded in the cab during every BTV firing except BTV #2, which was also the only BTV which did not have the backblast directed at the vehicle cab. The NIOSH short duration limit (140 mg/m³ for 30 minutes or less) for HCl was not exceeded during the two round ripple firings of BTVs 3 and 4, 7 and 8, and 9 and 10. BTV #1 and the two round ripple firing of BTV 5 and 6 exceeded both the OSHA and NIOSH limits.

(2) The 440 ppm CO level measured during the ripple firing of BTV #5 and #6 exceeds the NIOSH ceiling level of 200 ppm CO.

(3) The impulse noise levels and B-Durations measured in the vehicle cab during the firing of BTV #1 were approximately 2 dB above the allowable limits for personnel wearing single hearing protection, according to Figure 5 of MIL-STD-1474. The

UNCLASSIFIED

UNCLASSIFIED

MIL-STD-1474 criteria indicate that personnel wearing double hearing protection (both ear plugs and ear muffs) would be adequately protected from the BTV #1 noise levels. The firing of BTV #1 blew open the doors of the CROSSBOW cab; and the gunner's side door was blown open again by the ripple firing of BTV #5 and #6. The sound levels measured in the cab during the other BTV firings when the doors were not blown open, were within the MIL-STD-1474 limits for personnel wearing single hearing protection, such as a sound attenuating communications helmet.

(4) The 140 dB contour around the CROSSBOW, which is the area in which hearing protection must be worn, is determined by the BTV noise rather than by the machine gun noise, because a BTV is louder than the machine gun. The 140 dB contour is a circle around the CROSSBOW with a radius of 519 feet. In a situation in which the machine gun will be fired but no missiles will be fired, the 140 dB contour is a circle around the CROSSBOW with a radius of 288 feet.

c. The following analysis applies to the toxic gas and noise levels measured in the AVENGER during PFW firings.

(1) The CO, NO₂, and SO₂ levels measured were within the OSHA limits of 50 ppm, 5 ppm, and 5 ppm respectively.

(2) The impulse noise levels from firing the PFW are within the allowable levels specified by MIL-STD-1474 for persons wearing single ear protection, such as sound attenuating communications helmet.

d. The following analysis applies to the toxic gas and noise levels measured in the CROSSBOW during PFW firings.

(1) The NO₂ levels measured during PFW firing were below the OSHA limit of 5 ppm, which is a "ceiling" limit that applies regardless of the duration of the exposure. The OSHA limits for CO and SO₂ (50 ppm and 5 ppm respectively) are not "ceiling" limits but are time-weighted average (TWA) limits that apply to exposures of 8 hours per day, and higher levels are allowable for short exposures. The CO and SO₂ measured during PFW firings did not exceed short term limits for these gases. The NIOSH ceiling limit for CO is 200 ppm. A 30 minute SO₂ exposure limit of 100 ppm is specified by NIOSH.

(2) The impulse noise levels from firing the PFW are within the allowable levels specified by MIL-STD-1474 for persons wearing single ear protection, such as sound attenuating communications helmets.

UNCLASSIFIED

UNCLASSIFIED

e. The following are the classifications of the identified safety hazards in accordance with the guidance of MIL-STD-882B.

(1) The hazard from excessive HCl levels entering the AVENGER cab when the missile backblast is directed at the cab, is classified as "catastrophic" in severity and "probable" in frequency. The excessive CO levels that accompany the HCl from the missile exhaust are "marginal" in severity.

(2) The hazard from excessive HCl levels entering the CROSSBOW cab when the missile backblast is directed at the cab, is classified as "catastrophic" in severity and "probable" in frequency. The excessive CO levels that accompany the HCl from missile exhaust are "critical" in severity.

(3) Paragraph 5.8.1:3 of MIL-STD-1472 states that crewstation effective temperature should not exceed 85 degrees F (29 degrees C) for extended periods of time. At effective temperatures above 39 degrees C the work periods are limited to 1/2 hour. For work periods exceeding 1/2 hour, heat stress will result. The hazard of heat stress is classified as "critical" in severity. This hazard applies to both AVENGER and CROSSBOW crew stations, under circumstances where the system is deployed in an environment similar to the environment of the hot chamber, in which the dry bulb temperature was 140 degrees F (60 degrees C).

(4) The hazards associated with using the AVENGER radio antenna or antenna post as a handhold for entering the gunner's turret (namely that the antenna will bend and also can cause a RF burn or an electrical shock) is classified as "critical" in severity and "probable" in frequency. The hazard could be controlled by warning labels on the antenna post, but preferably the antenna should be relocated.

(5) The hazard associated with Avenger gunners receiving minor cuts while loading missiles or machine guns is classified as "marginal" in severity and "occasional" in frequency.

(6) The hazard of AVENGER gunners bumping against the ventilator inlets' sharp corners is classified as "marginal" in severity and "occasional" in frequency.

(7) The hazard of the CROSSBOW missile becoming damaged from striking a radio antenna is classified as "critical" in severity and "remote" in frequency.

(8) The hazard of the HMMWV filter capacitors sparking and damaging a battery is classified as "marginal" in severity and "remote" in frequency.

UNCLASSIFIED

APPENDIX C
PRELIMINARY TRAINING REQUIREMENTS ANALYSIS

PRELIMINARY TRAINING REQUIREMENTS ANALYSIS PEDESTAL MOUNTED STINGER

INTRODUCTION

A training requirements analysis was conducted using the principles outlined for TRADOC Systems Approach to Training (SAT) appropriately tailored to support the test and evaluation of the PMS candidate systems. The objective of the analysis was to determine if either of the PMS candidates will require training that cannot be provided within the training base available or planned for MOS 16S.

The procedures determined to be appropriate for this analysis was to conduct a task analysis, determine institutional training requirements, and determine unit training requirements.

TASK ANALYSIS

A task analysis was conducted beginning with mission tasks derived from functions demonstrated by each of the candidate systems, and tactical and doctrinal procedures identified for the test and evaluation.

Tasks were initially determined by a review of the individual operators manual for each candidate system, and observations of test missions and activities. All modes of operation; on-the-move, stationary, remote and MANPADS, were evaluated. When completed, the task list was given to each contractor for review and comment. Comments from contractor representatives were integrated into the task list.

Task analysis between the candidate systems did not yield significant levels of task difference. Although, there were differences in the number of functional tasks (tasks identified one indenture below the mission level), the differences are attributed to the manner in which the contractor chose to describe the operational activities. Minor differences in equipment/software that specifically influence tasks are shown in Table 1.

Explanation of these equipment/software differences also does not drive significant differences in the candidate systems. The Boeing system requires the operator initiate the built in test (BIT) with built in test equipment (BITE) during system functional checks and initialization. The LTV system operates a continuous BIT which does not require the operator initiate specific checks using this technology. The difference for the operator is the few seconds required to initiate the BITE on the Boeing system. The Primary Power Unit (PPU) on the LTV system is a typical power

generating unit that requires only a few steps to check out and place in operation. Power is supplied to the Boeing system through batteries charged by the vehicle. The difference between the optical sight on the Boeing system and the TV on the LTV system does not drive workload differences.

INSTITUTIONAL TRAINING REQUIREMENTS

Mission tasks identified in the initial task listing were further analyzed to identify tasks selected for training. To provide a basis for comparison, the tasks identified for training were then clustered into subject areas to match the current MANPADS institutional training course 043-16S10. These subject areas are shown in Table 2.

The two candidate systems were then compared to the current MANPADS course. Neither the number of tasks nor the instructional hours were found to be significantly different. There are only 18.5 hours difference in the MANPADS Course and the course requirement for the PMS candidates. Based on an anticipated 16 repetitions of the course during a year, 296 additional annual instructional manhours are required for the PMS. The 16 repetitions are the number of times the course must be repeated to replenish losses to the PMS forces at "steady state."

An instructor at the Army schools is available to perform his duties 1,728 hours each year. Based on this instructor availability, the additional requirement for the PMS candidates is .17 instructor, or no additional instructor requirement.

Although the MANPADS course is the footprint for evaluating the PMS candidate systems, it does not provide sufficient comparison for evaluating the additional task requirements. Restructuring of the subject areas for the PMS tasks selected for training allows for a better evaluation of the opportunity to adequately train personnel to operate the systems. Course 043-16P10 is the Air Defense training course for the Chaparral weapon system. The subject areas outline is shown in Table 3.

A quasi-program of instruction (QPOI) was developed from this analysis. A course summary identifying instructional requirements to teach all of the PMS tasks is shown in Table 4.

To provide a broader comparison to the PMS system, Table 5 identifies the instructional hours required, by subject areas, for the notional PMS course and three other Air Defense courses; Chaparral, Vulcan, and MANPADS.

Based on this preliminary analysis, there appear to be not significant problems in training the tasks required to operate either of the PMS candidate systems. The training course summary developed during this analysis is a reasonable assessment of the PMS institutional training requirements and is consistent with other USAADASCH and TRADOC training concepts.

UNIT TRAINING REQUIREMENTS

Unit Commanders will continue to be required to initially train 20% of the individual critical tasks and sustain proficiency in all of the tasks. To accomplish their job, commanders will have a series of products, programs and devices available similar to those listed in Table 6. Only embedded training devices represent a new focus on the unit training requirement. The exact application of the embedded training is a subject of separate analysis. All other approaches are consistent with current unit training programs.

**TABLE 1. EQUIPMENT DIFFERENCES
AFFECTING TASKS**

<u>BOEING SYSTEM</u>	<u>LTV SYSTEM</u>
Operator Initiated BITE	——
——	Primary Power Unit (PPU)
Optical Sight	TV

TABLE 2. SUBJECT AREAS FOR 16S COURSE

- o **Visual Aircraft Recognition**
- o **Introduction to 16S Training, Operation and Maintenance of the Stinger Weapon System, Training Devices, and Programming the IFF Interrogator**
- o **Operation of the Redeye Weapon System and Training Devices**
- o **Operation and Maintenance of the M151 1/4-Ton Truck, M416 Utility Trailer, and Communications Equipment**
- o **Command and Control**
- o **EOCCT/FTX**

TABLE 3. SUBJECT AREAS FOR PMS (NOTIONAL) COURSE

- o **Course Orientation**
- o **Visual Aircraft Recognition**
- o **Operation and Maintenance of the Carrier (HMMWV) Subsystem**
- o **Operation and Maintenance of the Communications Subsystem**
- o **Operation and Maintenance of the PMS Subsystem**
- o **Team Operations**
- o **EOCCT/FTX**

TABLE 4. COURSE SUMMARY

Annex	Subject Area	Academic Hours
Annex A	Course Orientation	2 (2C)
Annex B	Visual Aircraft Recognition	32 (30 PE3; 2 E2)
Annex C	Operation and Maintenance of the Carrier (HMMWV) Subsystem	30 (2C; 2D; 23 PE1; 3 E1)
Annex D	Operation and Maintenance of the Communications Subsystem	30 (3C; 5D; 15 PE1; 6 E1; 1 E3)
Annex E	Operation and Maintenance of the PMS Subsystem	106 (16.5C; 3D; 66.5 PE1; 8 E1; 2 E3)
Annex F	Team Operations	24.5 (3.5C; 3.5D; 17.5 PE1)
Annex G	EOCCT/FTX	47 (41 PE1; 6 E1)
Total Academic Hours		271.5

Class Size:	Maximum 36	Optimum 30	Minimum 24
-------------	---------------	---------------	---------------

Classes per year: 16

Class start: 3 week intervals

Class length: 7.5 weeks

Type instruction by hour:	C - 37
	D - 13.5
	PE1 - 163
	PE3 - 30
	E1 - 23
	E2 - 2
	E3 - 3
	<hr/>
	Total 271.5

Active/Passive Evaluation Index:

Active
71%

Passive
19%

Examination
10%

**TABLE 5. COMPARISON OF ACADEMIC COURSE REQUIREMENTS
AIR DEFENSE SYSTEMS**

Subject Area	*XXX- 16X10	043- 16P10	041- 16R10	043- 16S10
Course Orientation	2	2	6	15
Visual Aircraft Recognition	32	32	32	32
Operation and Maintenance of the Carrier (Vehicle) Subsystem	30	65	35	41
Operation and Maintenance of the Communications Subsystem	30	43	23	19
Operation and Maintenance of the Fire Control/ Weapons Subsystem	106	91	94	75
Team Operations	24.5	18	9	24
EOCCT/FTX TOTAL	<u>47</u> 271.5	<u>45</u> 296	<u>47</u> 246	<u>47</u> 253

TABLE 6. UNIT TRAINING PRODUCTS, PROGRAMS AND DEVICES

Soldier's Manuals

Job Books

STX/Drills

AMTP

Common Troop Proficiency Trainer

2D/3D Integrated Video Disc

Dummy/Smart Missile Simulators

Launch Simulator

1/5 Scale Targets

Combat Tables

Multi-purpose Range Complex

Embedded Training Devices

PROGRAM OF INSTRUCTION COMPARATIVE ANALYSIS
 MAINTENANCE CREW (040-10010) to FMS Tactical Crew XII-10410

Subject Area	MAINTENANCE Crew Task #	Task Description	Learning Comp Task #	Task Description	LTV Comp Task #	Task Description
Visual Aircraft Recognition	441-040-1040	Visually Identify Threat and Friendly Aircraft	441-000-1040	Visually Identify Threat and Friendly Aircraft	441-100-1040	Visually Identify Threat and Friendly Aircraft
		<p>Introduction to ICS Training, Operation and Maintenance of the Radar Display System, Training Devices, and Programming the IFF Interrogator</p>	441-000-1027	Perform Observer Functions During Target Engagement with the FMS Fire Unit	441-100-1027	Perform Observer Functions During Target Engagement with the FMS Fire Unit
			441-000-1081	Prepare the FMS Fire Unit for a Tactical Air Defense Mission	441-100-1081	Prepare the FMS Fire Unit for a Tactical Air Defense Mission
			441-000-1005	Load/Reload Missiles on the FMS Fire Unit	441-100-1005	Load/Reload Missiles on the FMS Fire Unit
			441-000-1041	Unload Missiles from the FMS Fire Unit	441-100-1041	Unload Missiles from the FMS Fire Unit
			441-000-1042	Load/Reload the .50 Caliber Machinegun	441-100-1042	Load/Reload the .50 Caliber Machinegun
			441-000-1043	Unload the .50 Caliber Machinegun	441-100-1043	Unload the .50 Caliber Machinegun
			441-000-1006	Perform Search Procedures Using the FMS FLIR Saboteur	441-100-1006	Perform Search Procedures Using the FMS FLIR Saboteur
			441-000-1045	Determine Target Range	441-100-1045	Determine Target Range
			441-000-1008	Perform Functional Checks on the FMS Using BITR	441-100-1044	Perform Search Procedures Using the TV Saboteur
			441-000-2023	Install/Test the Identification Friend or Foe (IFF) Interrogator AH/PFA-3 on the FMS Fire Unit	441-100-2017	Start the Power Distribution Unit (PDU) on the FMS Fire Unit
			441-000-3002	Perform Before and During Operations Preventive Maintenance Checks and Services (PMCS) on the IFF Programmer/Battery Charger	441-100-3002	Perform Before and During Operations Preventive Maintenance Checks and Services (PMCS) on the IFF Programmer/Battery Charger

PROGRAM OF INSTRUCTIONS COMPARATIVE ANALYSIS
 HARPAS Crew (043-10810) to PMS National Crew III-10110

Subject Area	HARPAS Task #	Task Description	Boeing Crew Task #	Task Description	LTV Crew Task #	Task Description
	441-046-1000	Perform Emergency Procedures for Pod Maladies			591-10v-1302	Prepare a DA Form 2404
	441-046-3003	Perform Before and During Operations Preventive Maintenance Checks and Services (PMCS) on the I77 Interrogator			441-10v-1046	Manually Range Target using TV/FLIR Reticle
	441-046-3003	Charge the I77 Interrogator Batteries	441-100-3003	Charge the I77 Interrogator Batteries	441-10v-3003	Charge the I77 Interrogator Batteries
	441-046-3004	Program the I77 Interrogator	441-100-3004	Program the I77 Interrogator	441-10v-3004	Program the I77 Interrogator
	441-047-1004	Engage Targets with the Stinger Weapon	441-100-1000	Perform Gunner Duties During Target Engagement with the PMS Fire Unit	441-10v-1030	Perform Gunner Duties During Target Engagement with the PMS Fire Unit
	441-047-1007	Convert a Stinger Missile Bound to a Weapon Bound	441-100-1007	Convert from PMS Configuration to HARPAS Configuration	441-10v-1007	Convert from PMS Configuration to HARPAS Configuration
	441-047-1008	Perform Abbreviated Weapon Checks, Stinger Weapon	441-100-1008	Perform Abbreviated Weapon Checks, Stinger Weapon (HARPAS)	441-10v-1008	Perform Abbreviated Weapon Checks, Stinger Weapon (HARPAS)
	441-047-1009	Perform Preventive Maintenance Checks and Services (PMCS), Stinger Weapon	441-100-1009	Perform Preventive Maintenance Checks and Services (PMCS) on the PMS Fire Unit	441-10v-1009	Perform Preventive Maintenance Checks and Services (PMCS) on the PMS Fire Unit
	441-047-1012	Perform Operator Corrective Maintenance Procedures, Stinger Weapon	441-100-1112	Perform Operator Corrective Maintenance Procedures on the PMS Fire Unit	441-10v-1112	Perform Operator Corrective Maintenance Procedures on the PMS Fire Unit
	441-047-1026	Ready the Stinger Weapon for Firing	441-100-1026	Ready the Stinger Weapon for Firing (HARPAS)	441-10v-1026	Ready the Stinger Weapon for Firing (HARPAS)
	441-047-1027	Operate Stinger Tracking Head Trainer (THT)	441-100-1127	Operate Stinger Tracking Head Trainer (THT) (HARPAS)	441-10v-1127	Operate Stinger Tracking Head Trainer (THT) (HARPAS)

PROGRAM OF INSTRUCTION COMPARATIVE ANALYSIS
 MAPAS Cms (042-10310) to MS Technical Cms XII-10310

Subject Area	MAPAS Task #	Task Description	Being Comp Task #	Task Description	LTV Comp Task #	Task Description
	441-007-1001	Perform Operational Check Procedures, Stinger Tracking Head Trainer (THT)	441-000-1101	Perform Operational Check Procedures, Stinger Tracking Head Trainer (THT)	441-100-1101	Perform Operational Check Procedures, Stinger Tracking Head Trainer (THT)
	441-007-1002	Perform Preventive Maintenance Checks and Services (PMCS), M134 Training Set	441-000-1002	Perform Preventive Maintenance Checks and Services (PMCS), M134 Training Set	441-100-1002	Perform Preventive Maintenance Checks and Services (PMCS), M134 Training Set
	441-007-1003	Perform Emergency Procedures for Bagfire/Miscfire, Stinger Weapon	441-000-1003	Respond to Missile Bagfires, Miscfires, and Beds	441-100-1003	Respond to Missile Bagfires, Miscfires, and Beds
Operation of the Redeye Weapon System and Training Services	441-070-1043	Inspect the Seven Critical Points, Redeye Weapon				
	441-070-1043	Recover Redeye Weapon to Prevent Empty Use				
	441-070-1044	Perform Preventive Maintenance Checks and Services (PMCS), Redeye Weapon				
	441-070-1045	Engage Targets with the Redeye Weapon				
	441-070-1047	Charge the M49 Tracking Head Trainer (THT) Batteries				
	441-070-1048	Operate Redeye Tracking Head Trainer (THT)				
	441-070-1051	Perform Preventive Maintenance Checks and Services (PMCS), M76 Training Set				
	441-070-1052	Perform Emergency Procedures on the Redeye Weapon				
	441-070-2010	Execute Fire Control, Redeye Gunner				

PROGRAM OF INSTRUCTION COMPARATIVE ANALYSIS
NAFAMS Cms (043-10010) to FMS National Cms XII-14310

Subject Area

Operation and Maintenance of
the FM 1 1/2-Ton Truck, FM 16
Fueling System, and
Communications Equipment

NAFAMS Task # Task Description FMS Task # Task Description

NAFAMS Task #	Task Description	FMS Task #	Task Description
441-100-1001	Perform Functional Checks on the FM 1 1/2-Ton Truck, FM 16 Fueling System, and Communications Equipment (Daily)	441-100-1001	Perform Functional Checks on the FM 1 1/2-Ton Truck, FM 16 Fueling System, and Communications Equipment (Daily)
441-100-1009	Replace Communications Equipment on the FM 1 1/2-Ton Truck	441-100-1009	Replace Communications Equipment on the FM 1 1/2-Ton Truck
441-100-1017	Start the FM 1 1/2-Ton Truck Auxiliary Power Source	441-100-1017	Start the FM 1 1/2-Ton Truck Auxiliary Power Source
441-100-1016	Troubleshoot the FM 1 1/2-Ton Truck	441-100-1016	Troubleshoot the FM 1 1/2-Ton Truck
441-100-1011	Perform Before Operations Preventive Maintenance Checks and Services (PMCS) on the FM 1 1/2-Ton Truck	441-100-1011	Perform Before Operations Preventive Maintenance Checks and Services (PMCS) on the FM 1 1/2-Ton Truck
441-100-1023	Operate the FM 1 1/2-Ton Truck Communications Equipment in a Tactical Environment	441-100-1023	Operate the FM 1 1/2-Ton Truck Communications Equipment in a Tactical Environment
441-100-1015	Perform Monthly Preventive Maintenance Checks and Services (PMCS) on the FM 1 1/2-Ton Truck	441-100-1015	Perform Monthly Preventive Maintenance Checks and Services (PMCS) on the FM 1 1/2-Ton Truck
441-100-1014	Perform Weekly Preventive Maintenance Checks and Services (PMCS) on the FM 1 1/2-Ton Truck	441-100-1014	Perform Weekly Preventive Maintenance Checks and Services (PMCS) on the FM 1 1/2-Ton Truck
441-100-1013	Perform After Operations Preventive Maintenance Checks and Services (PMCS) on the FM 1 1/2-Ton Truck	441-100-1013	Perform After Operations Preventive Maintenance Checks and Services (PMCS) on the FM 1 1/2-Ton Truck
441-100-1012	Perform During Operations Preventive Maintenance Checks and Services (PMCS) on the FM 1 1/2-Ton Truck	441-100-1012	Perform During Operations Preventive Maintenance Checks and Services (PMCS) on the FM 1 1/2-Ton Truck

112-007-0043 Prepare/Operate FM Radio Set

112-008-1015 Install and Operate Field Telephone

PROGRAM OF INSPECTION COMPARATIVE ANALYSIS
NAEPAD Crs (043-10210) to FB National Crs III-10210

Subject Area	NAEPAD Tank #	Tank Description	Reading Comp Tank #	Tank Description	LTV Comp Tank #	Tank Description
	119-000-3017	Perform Operator's POC/Exercise Checks on Telephone Set 74-312/77				
	351-721-1027	Operate Vehicle with Manual Transmissions	441-100-1010	Drive the M20V	441-100-1010	Drive the M20V
Ground and Control	441-007-1011	Destroy Stinger Weapon to Prevent Enemy Use	441-100-1111	Destroy the FB Fire Unit to Prevent Enemy Use	441-100-1111	Destroy the FB Fire Unit to Prevent Enemy Use
	441-007-3015	Selection of NAEPAD Firing Positions	441-100-2015	Selection of NAEPAD Firing Positions	441-100-2015	Selection of NAEPAD Firing Positions
	441-007-3016	Exercise Fire Control, Stinger Gunner	441-100-2016	Exercise Fire Control, Stinger Gunner (NAEPAD)	441-100-2016	Exercise Fire Control, Stinger Gunner (NAEPAD)
	071-320-1014	Locate on Unknown Point on a Map or Ground by Intersection	071-100-1014	Locate on Unknown Point on a Map or Ground by Intersection	071-100-1014	Locate on Unknown Point on a Map or Ground by Intersection
	071-320-1015	Locate on Unknown Point on a Map or Ground by Intersection	071-100-1015	Locate on Unknown Point on a Map or Ground by Intersection	071-100-1015	Locate on Unknown Point on a Map or Ground by Intersection

02/20/77

FOUNDATIONS FOR AN OPERATOR WORKLOAD (OWL)

ASSESSMENT PROGRAM FOR THE ARMY

F. A. Glenn, III
Alvah C. Bittner, Jr.
Walter W. Wierwille
Helene P. Iavecchia
Robert J. Wherry, Jr.
Paul M. Linton



U.S. Army Research Institute
for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria VA 22333

This working paper is an unofficial document intended for limited distribution to obtain comments. The views, opinions, and/or findings contained in this document are those of the author(s) and should not be construed as the official position of ARI or as an official Department of the Army position, policy, or decision, unless so designated by other official documentation.

Draft Technical Report 2075-1

FOUNDATIONS FOR AN
OPERATOR WORKLOAD (OWL)
ASSESSMENT PROGRAM FOR THE ARMY

December 1986

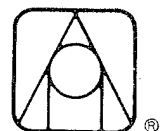
Prepared by:

Floyd A. Glenn III, Ph.D.
Alvah C. Bittner, Jr., Ph.D.
Walter W. Wierwille, Ph.D.
Helene P. Iavecchia
Robert J. Wherry, Jr., Ph.D.
Paul Linton

Submitted to:

Dr. Richard Christ
Army Research Institute Field Unit
P.O. Box 6057
Ft. Bliss, TX 79906-0057

Task Report
Contract No. MDA903-86-C-0384
Item No. 0002AD



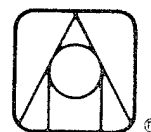
EXECUTIVE SUMMARY

This report is the first of a series describing a program for the development and validation of a methodology for estimation and evaluation of operator workload (OWL) in Army systems. It presents the results of Task 1 of Analytics' contract with the Army Research Institute (ARI) to "Identify OWL Concepts and Measures, and Develop a Working Model of OWL". Included are the results of component subtasks: (1.1) Provide Operational Definitions, (1.2) Identify Existing Measures, and (1.3) Develop Workload Model. The overall purpose of this report is to establish a foundation for the OWL program and to identify relevant prior research and techniques.

The principal elements of the foundation of the OWL program are presented. OWL is defined as the relative limitation in the human's ability to do work, rather than as some single- or multi-dimensional quantity which might be directly measurable. An OWL performance model is described which relates OWL to all other key aspects of human and system performance (e.g., physiology, subjective state, system design, mission requirements, etc.). This model provides a conceptual tool to support subsequent efforts in the OWL program. A review is presented of currently available tools and measures for the assessment of OWL along with a comprehensive OWL taxonomy broken into the major categories of analytical methods and empirical methods. Strategies and tools for the management of documents and information pertinent to the OWL program are described. Finally, plans for the subsequent four tasks in the OWL program are presented. These include:

Task 2 -- Identify Army Requirements Regarding OWL, Select Specific Army Systems to Analyze, and Provide Outline of OWL Handbooks;

Task 3 -- Evaluate Measures of OWL and Develop Validation Plan;



Task 4 -- Validate Measures and Conduct OWL Analysis of the
Prototype Systems;

Task 5 -- Prepare Pamphlet, Handbooks, Survey and Report.

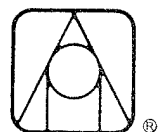


TABLE OF CONTENTS

1.	INTRODUCTION	
1.1	Overview of OWL Program	1-1
1.2	Background	1-2
1.3	Organization of Report	1-5
2.	DEFINITIONS	
3.	OWL PERFORMANCE MODEL	
3.1	Objectives of Model	3-1
3.2	Performance Model Description	3-2
3.2.1	Task Demands	3-2
3.2.2	Human Operator	3-2
4.	REVIEW OF OWL MEASURES AND TECHNIQUES	
4.1	Taxonomy	4-1
4.2	Analytical Techniques	4-4
4.2.1	Expert Opinion	4-4
4.2.2	Comparison	4-5
4.2.3	Simulation Models	4-6
4.2.4	Math Models	4-7
4.2.5	Task Analytic Methods	4-9
4.3	Empirical Techniques	4-10
4.3.1	Operator Opinion Measures	4-11
4.3.2	Primary Task Measures	4-14
4.3.3	Secondary Task Measures	4-16
4.3.4	Physiological Measures	4-20
5.	OWL MANAGEMENT INFORMATION SYSTEM (OWLMIS)	
5.1	OWL Information Data Base (OWLIDB)	5-1
5.2	OWL Information Analysis System (OWLIAS)	5-2

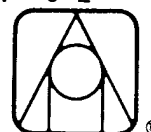
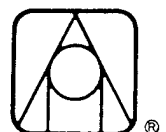


TABLE OF CONTENTS (Continued)

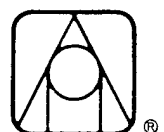
5.3	OWL Library (OWLLIB)	5-2
5.3.1	Formal Sources	5-3
5.3.2	Informal Sources	5-3
5.3.3	Current Status	5-4
6.	SUMMARY AND PLANS	
6.1	Summary	6-1
6.2	Task 2 Plans	6-1
6.2.1	Formal Army Requirements	6-2
6.2.2	Assessment of User Needs	6-2
6.2.3	Outline of Handbooks	6-5
6.2.4	Selection of Specific Systems	6-6
6.3	Plans for Subsequent Tasks	6-7

REFERENCES



FIGURES

1-1	OWL Matching Model	1-4
3-1	OWL Performance Model	3-3
3-2	A Two-Stage Physiological State Model	3-5
3-3	Individual Differences in Workload Response as Illustrated by Changes in Intercondition Correlation	3-8
4-1	Expandable Taxonomy of Workload Assessment Techniques	4-3



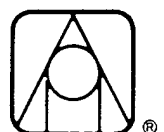
1. INTRODUCTION

This report is the first of a series describing a program for the development and validation of a methodology for estimation and evaluation of operator workload (OWL) in Army systems. It presents the results of Task 1 of Analytics' contract with the Army Research Institute (ARI) to "Identify OWL Concepts and Measures and Develop a Working Model of OWL". Included are the results of component subtasks: (1.1) Provide Operational Definitions, (1.2) Identify Existing Measures, and (1.3) Develop Workload Model. The overall purpose of this report is to establish a foundation for the OWL program and to identify relevant prior research and techniques.

1.1 OVERVIEW OF OWL PROGRAM

The focus of the OWL program is on the practical problem of determining what the Army can and should do to assure that systems can be adequately operated by prospective Army personnel. Over more than two decades, a considerable body of research has addressed various facets of the workload question, and much of this is relevant to Army concerns. However, the OWL program is not simply an attempt to consolidate previous research efforts. Its intent, rather, is to select and integrate the most useful elements from these efforts for aiding developers and evaluators of combat systems. It is important for the Army to make the best possible use of existing workload assessment technology to assure the acceptable performance of Army systems.

There are a substantial number of workload measurement techniques which are suitable for measuring some factor that impacts system performance. Each technique has significant characteristics including requirements, costs of



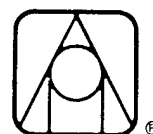
implementation, and reliability of results. The principal output of the program will be a systematic OWL estimation and evaluation methodology. This methodology will be documented in Army handbooks, which will guide users in OWL management throughout all phases of the system development cycle. The OWL methodology and its component measurement techniques will also be empirically validated in this program, primarily through application to several Army systems that are in different stages of development. The result of the OWL program will be validated methods for OWL management along with detailed user guidance presented in Army handbooks.

1.2 BACKGROUND

The roots of OWL research go back at least to the early industrial engineering work of Frederick Taylor (e.g., Taylor, 1912) and Frank and Lillian Gilbreth (e.g., Gilbreth, 1911, and Gilbreth and Gilbreth, 1917). This work, however, focused on the measurement of human performance of simple, repetitive, manual tasks, especially tasks involved in assembly line production. These time-and-motion study methods continue to be applied to manual work environments in industry but have not been successfully expanded to more complex situations involving cognitive work and more open-ended tasks.

The contemporary period of OWL research, considering both manual and cognitive work, appears to have begun around 1960. The earliest usage we have found of the term "workload" was in a paper by Knowles (1963), though we found a slightly earlier reference to the related term "operator loading" (Siegel and Wolf, 1961). Since that time, topics concerned with OWL have been popular subjects of research, and a substantial literature has accumulated in this area. This research has addressed a broad range of issues including:

- Identification and validation of OWL measures;
- Impact of a variety of factors such as individual differences, training, and environmental conditions on OWL;
- Development of models for OWL and related aspects of human performance.



Workload analysis is currently mandated for all major phases of development of Department of Defense (DoD) systems by Military Specification MIL-H-46855B (Human Engineering Requirements for Military Systems, Equipment and Facilities). Section 3.2.1.3.3 of that MIL-SPEC states succinctly that "Individual and crew workload analysis shall be performed and compared with performance criteria." However, the document does not further define what is considered to constitute a workload analysis, though other sections do call for a variety of analyses that appear to relate to workload. This ambiguity accounts for a high degree of variability in the manner in which system developers address the OWL issue. In some cases, OWL may be perceived as a key problem, as in the Army's LHX scout/attack helicopter, and considerable development resources may be allocated to OWL assessment. In other cases, where OWL is initially perceived as a minor issue, it may be given only cursory attention with subjective paper-and-pencil techniques. The meaning of the mandate for OWL analysis is consequently uncertain and appropriate assessments of workload are not insured.

Because of the recognized importance of OWL in Army systems and the large number and diversity of techniques for measuring OWL, it is important for the Army to establish an effective, standardized methodology for OWL assessment in developing systems. Hence, a methodology is required to aid in determining how OWL should be assessed for each Army system according to system characteristics, system/mission requirements, and system development phase. We envision this methodology as an explicit matching model, illustrated in Figure 1-1, which derives an OWL assessment battery from specific consideration of OWL measure characteristics, system characteristics and requirements, and an OWL performance model. The OWL performance model, discussed in detail in Section 3, provides a vehicle for identifying pertinent interrelations of OWL measures, such as indications of how measures might complement or supplement each other. The matching model may take a variety of different forms, depending largely on the results of the evaluation of OWL measures in Task 3 of this project. Possibilities for the matching model range from a set of narrative guidelines to an automated expert system. One expert system that performs this type of function has recently been developed by the Army and NASA (Casper, et al., 1986) and



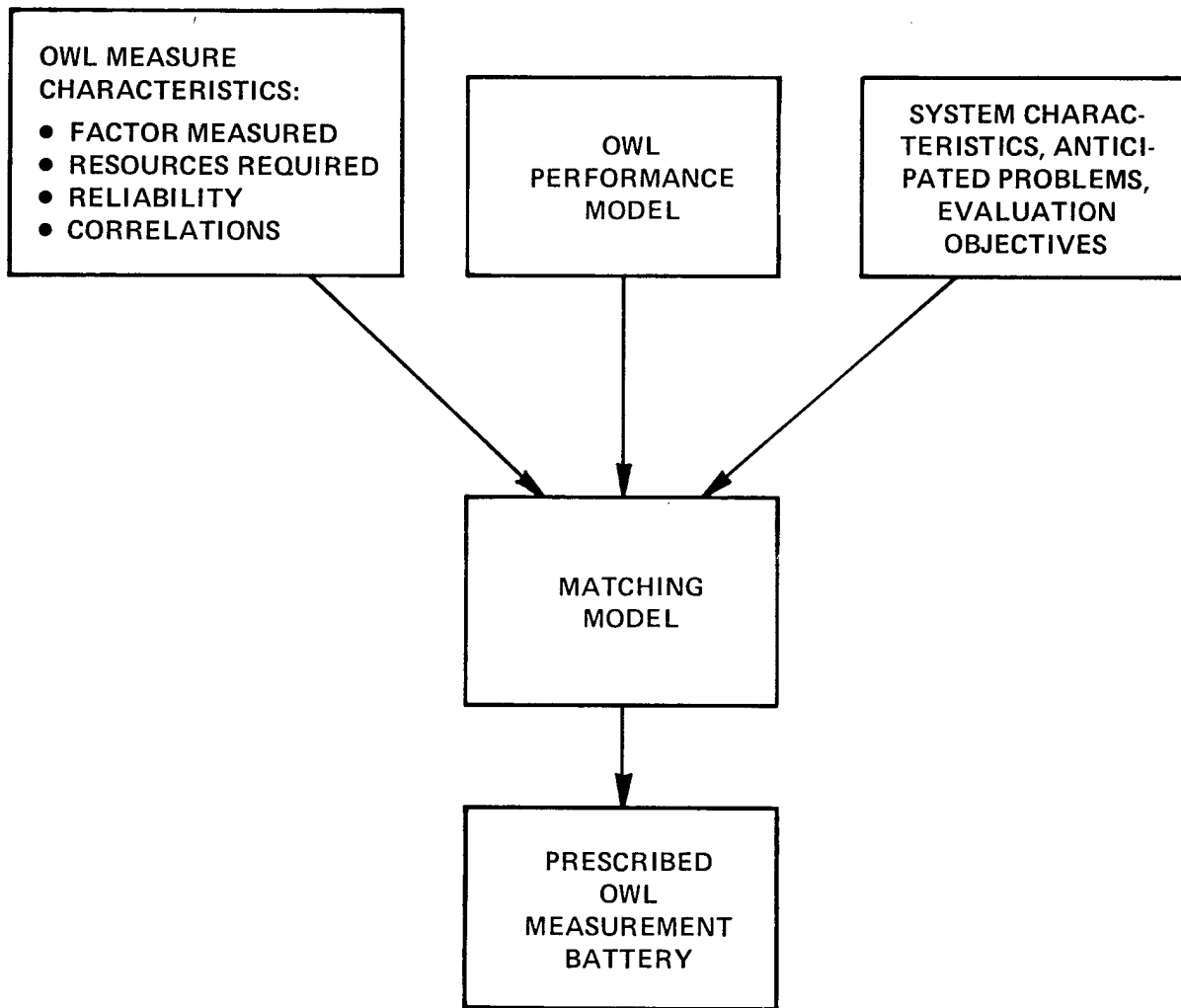


Figure 1-1. OWL Matching Model

will be considered for applicability for the OWL methodology. The matching model will be definitized and documented in the course of a subsequent effort (Task 3 of the OWL program).

1.3 ORGANIZATION OF REPORT

The body of this report presents further details of the results of Task 1. Section 2 addresses the problem of establishing a working definition of OWL and relating it to alternative definitions that are commonly used. The OWL definition is elaborated in Section 3 through development of an OWL performance model which offers a comprehensive concept of how OWL is related to other aspects of human-system performance. A brief review of existing OWL assessment techniques is presented in Section 4. Section 5 then describes procedures that will be used throughout this project to collect and manage data on OWL assessment techniques and related OWL literature. Finally, Section 6 provides an overview of plans for the remainder of the OWL project.



2. DEFINITIONS

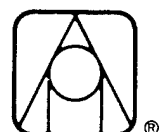
Before recommending a working definition of workload, it is instructive to consider what we want it to be. We can, of course, define the word to mean whatever we wish, though we must be careful to assure that the result will provide useful reference to some real process or condition. A common, idealized view of the term "operator workload" is represented in the following:

Naive Definition: Operator workload refers to a single dimensional measure of the aggregated load imposed on the operator by the job which he/she must perform. As the load is increased, acceptable performance is obtained until some load threshold is reached; above the threshold, errors and failures in performance become increasingly likely and severe.

This definition presumes some kind of generic work-performing resource or capacity which sustains/accomplishes the aggregated load. It is difficult to imagine a common capacity that supports all physical and cognitive work, though intense work of all types does tend to produce a common fatigue which is countered by a common type of rest. But fatigue is not the same thing as workload; fatigue is one of the results of sustained workload. Perhaps rather than a single workload dimension, some few dimensions might be identified with an associated multi-dimensional threshold for performance.

Most OWL research has been based, at least implicitly, on some more narrow and more operationally oriented definition. Some of the disparate views include identification of OWL as:

- The load that is imposed on the operator (i.e., objective task characteristics such as quantity of information to process and



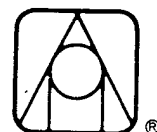
time allowed for processing) without regard to its effect on the operator or the performance;

- The subjective or physiological condition associated with a specified work effort (e.g., the affective burden placed on the operator by the work);
- The ability of the operator to perform other simultaneous tasks along with the primary task (i.e., residual capacity); and
- The functional relationship between primary task performance and the various parameters that impact that performance.

We believe that it is appropriate, for the purpose of operational workload evaluation, to treat all views of workload as viable. This belief follows from a functional perspective of how each workload measure conforms with differing constraints and objectives of an evaluation effort. The issue of any evaluation is not whether we can tap some "true" indicator of workload. Rather, the issue is whether we can determine if a man-machine system will perform at or above criterion levels under all specified operational conditions. Workload is only of interest, in the present context, insofar as it is related to system performance. The relationship may be fairly remote in some cases. High workload stress might result in low personnel retention in a certain MOS and a subsequent inability to man a weapons system adequately. In other cases the relationship will be direct; an overloaded operator will not be able to accomplish the prescribed mission in the available time.

The goal in the present case is to establish a definition which spans the considerations of all system development issues over all phases of the Army system life cycle. While a diverse collection of measures is required to address all issues and phases, a single definition is warranted to point to the common focus of these measures. For the purpose of developing a practical OWL assessment methodology for the Army, the appropriate OWL definition should be functionally oriented so as to indicate the operational significance of the OWL measures. We propose the following working definition:

Working Definition: OWL is a representation of a human operator's relative limitation in the capability to do work.



As an indicator of "relative limitation", OWL focuses on conditions under which unacceptable system performance would occur. OWL is typically assessed not to tell us how much work an operator is accomplishing, but rather to tell us how close the operator is to being unable to accomplish required work. We assume that, in this context, we are not interested in just an absolute measure of how much work the operator accomplishes. Rather, the interest is in determining how the work actually performed compares to both the mission requirement and the operator's ultimate capability to do work. It is important to recognize that many types of work and many types of limitation are subsumed in this definition. The common issue addressed by OWL, from the perspective of the system developer, cuts across all these distinctions. The issue is whether or not the system operators will be able to perform all required tasks adequately under specified operational conditions. OWL is only of interest to the system developer to the extent that it influences system performance.

The above working definition of OWL, it is important to note, implies that OWL is intrinsically unobservable. As a "relative limitation in capability," OWL is considered as a potential for unsatisfactory performance rather than an observable aspect of performance. The OWL concept thus falls in the category of hypothetical construct -- a theoretical process that is invoked to explain an observed pattern of behavior, with no direct evidence for the existence of the process. Primary and secondary task performance, physiological measures, and subjective reports may provide evidence of the operation of some OWL-related process, but they are not direct measures of OWL. An analytic model could hypothesize an explicit, measurable OWL process (like the network of workload capacity reservoirs that we describe in Section 3); however, an analytic model is just a formalization of the OWL construct, not a demonstration of its existence. For the present program, however, it is not critical to resolve such questions as whether or not there is a real physiological process which underlies all OWL effects on performance; it is critical to develop reliable procedures for estimating and evaluating those performance effects.

The types of work distinguished by OWL techniques can be effectively partitioned into perception, cognition, communication, and motor processes.



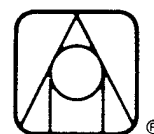
This follows the scheme adopted by Wierwille and his colleagues in their comprehensive study of workload measures for aircraft crewmembers (Wierwille et al., 1979). This human operator task taxonomy (called Universal Operator Behaviors) was presented earlier by Berliner et al., (1964). It is not critical that the distinctions between these categories be absolutely clean, as they are not, as long as the classification is useful for the OWL methodology. The value of the classification of work is to aid in matching OWL measures to system performance issues. Whereas alternative schemes could draw finer distinctions among motor processes (e.g., explosive force, sustained force, perceptual-motor coordination), the principal focus of the desired methodology is on mental workload. Hence, a scheme that emphasizes major classes of mental work is particularly appropriate (i.e., perception, cognition, communication).

Many types of performance limitations could be identified for each of the work categories. Some limitations, however, are best addressed outside the context of OWL; these include the data-limited processes which may be distinguished from resource-limited processes (cf., Norman and Bobrow, 1975). Data limitations occur when a process is constrained by its input data (e.g., sensory stimuli below threshold or insufficient information to solve a problem) rather than by the human's information processing capability. Though data limitations may affect workload, these are best addressed as interface design issues separate from OWL (e.g., assuring that the right information is displayed and is easily discriminable). Resource-limited processes can exhibit various types of performance degradation when the applicable limit is reached; performance may degrade gradually, intermittently, or catastrophically. Of course, the form and significance of any performance degradation is largely defined by the performance measures and criteria that are employed.

Another important distinction in the definition of OWL is that between empirical and predicted measures. Empirical measures are obtained from the performance of actual operators using a real or simulated system. Predicted measures are obtained through analysis, modeling, and/or estimation without empirical measurement of operators and systems of interest. Empirical measures are only feasible when a real or simulated system is available, along with



trained operators. Predicted measures, on the other hand, are the natural choice at early stages in system design or when systems or simulators are not available or feasible to use. Predicted measures based on detailed performance simulation have also been shown to be useful for operational systems because of the detailed diagnostic information they can provide and because of the ease with which many parameters can be manipulated (Lane, et al., 1979).

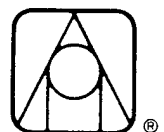


3. OWL PERFORMANCE MODEL

Considering its status as a hypothetical construct and its involvement with many other facets of human performance (e.g., attention, fatigue, stress, physiology, etc.), the OWL performance model must be developed to support many subsequent efforts in the OWL program. Most immediately, such a model is needed in order to provide a detailed definition of the scope of our interests in studying OWL and in identifying the Army's OWL-related needs. The model can also serve as a guide and as an information consolidation vehicle for our review and evaluation of OWL measures. As indicated in Figure 1-1, the OWL performance model will have a prominent role in the OWL matching model in order to provide detailed advice for operational workload assessment problems. It is even conceivable that the performance model might be eventually definitized to a sufficient degree that it could be incorporated as part of a simulation system such as the Human Operator Simulator (HOS, Harris, et al., 1986) or the Systems Analysis of Integrated Networks of Tasks (SAINT, Seifert and Chubb, 1978) and so used directly to assess OWL. The OWL performance model which we present here is intended as an evolutionary tool which we will apply, reexamine, and refine throughout the course of the OWL program.

3.1. OBJECTIVES OF MODEL

The OWL performance model is a framework for the understanding of the interplay of the diverse task and operator factors that influence workload. As such, one primary objective in its formulation was to develop a comprehensive structure for organizing and interpreting the extensive literature of OWL. Hopefully, this would provide for a method for integrating measures algorithmically as has been successfully accomplished in the domain of physical manual handling workload (cf., Karwowski and Ayoub, 1984). The organization and



interpretation of the OWL literature, of course, is aimed at the problem faced by designers and evaluators to efficiently and appropriately assess OWL during the various phases of the system acquisition cycle. Another primary objective for this formulation, because of the interest of systems designers and evaluators, was a careful delineation of the sources and nature of individual differences in OWL. The overall objective of the model is to provide a framework for the selection, use, and interpretation of OWL measures for analysis of combat.

3.2. PERFORMANCE MODEL DESCRIPTION

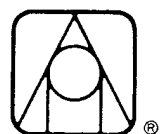
Figure 3-1 illustrates this model and distinguishes factors associated with task demands and the human operator. These are discussed in turn in the following.

3.2.1 Task Demands

The task demands which define the operator's mission are jointly determined by factors largely outside the operator. These include the system design, input load (e.g., information input rate imposed), environmental factors (e.g., weather and vibration), the operator's understanding of mission requirements, as well as social/cultural factors. In the OWL program, our interest in these factors is driven by the consideration of how they impact human performance; each of these factors is best examined in terms of how it influences each detailed aspect of human operator performance.

3.2.2 Human Operator

The human operator is viewed in the model as comprised of three interrelated facets -- physiological state, subjective state, and performance resources (cf., Fig. 3-1). From this view, performance resources refer to the operator's "potential" higher-level behavioral components which interact to accomplish task elements (e.g., knowledges, abilities, skills, strategies, strength, motivation, etc.). As will be indicated in the following, expressions of these "potential" resources are moderated by physiological and subjective



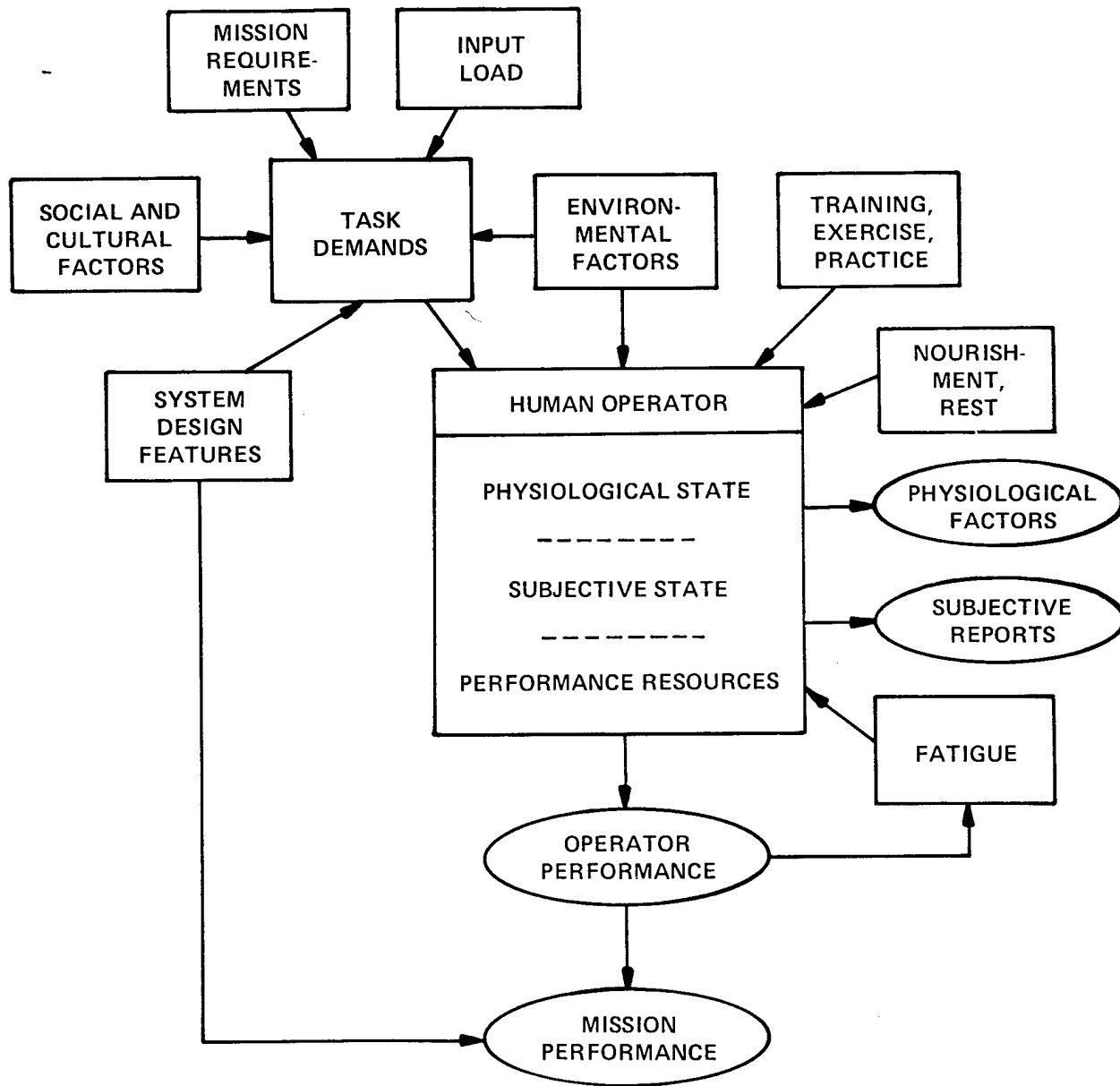


Figure 3-1. OWL Performance Model



status. Consequently, assessment or estimation of these "potentials" requires minimization or control for these status effects. Performance resources are a primary source of individual differences in workload and workload capacity.

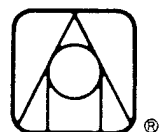
Moderating the utilization of the performance resources is the operator's physiological state. As suggested by Figure 3-1, this state is directly affected by environmental factors, shaped in the long term through training, exercise, and practice as well as impacted in the short term by such factors as nourishment and rest. We believe, as a working concept, that it is useful to consider (1) the physiological state to determine the "operator's readiness-to-respond" and (2) that this readiness is the end result of a cascade of multiple reservoirs driven by physiological resources for responding. Figure 3-2 illustrates this concept for a two-stage reservoir model where rest and nourishment provide the inputs to a long-term capacity for responding which, in turn, replenishes the short-term reservoir. As may be seen, the replenishment of the short-term reservoir is dependent on: (a) the state of depletion of the its reservoir and the mechanism of its output flow (faucet); and (b) the level of the long-term reservoir and its output mechanism. The output of the short-term reservoir at each moment of time represents the momentary amount of effort being expended. Simulation of the time-course of the momentary output of this two-stage reservoir model appears consistent with our experience with individual differences in physiological responses, and with operator performance. This model is analogous to one for manual handling where the ability to perform may be bounded both by short-term (e.g., biomechanical) and longer-term (e.g., aerobic) capabilities (i.e., Karwowski and Ayoub, 1984). A similar reservoir model has also been invoked by Kahneman (1973) to explain the manner in which simultaneous performance of two tasks may produce mutual interference (if the two tasks draw on the same reservoir) or not (if they draw on different reservoirs). More interestingly, the model both suggests the roles of short- and long-term response systems in OWL as well as pointing to the need for considering their assessment. Their assessment, of course, requires measures of physiological factors as indicated in Figure 3-1. Physiological factors are a source of individual differences in workload capacity which are currently approached from the framework of a two-stage physiological reservoir model.



Also moderating the utilization of the performance resources is the operator's subjective state. This state, in addition to being influenced by (and influencing) the actual physiological state, is substantially determined from "perceptions" of the task demands and physiological state. Changing such perceptions, by cognitive-behavioral and other approaches, frequently produces dramatic physiological stress indicator reductions as well as freeing-up performance resources (cf., Mostofsky & Piedmont, 1985). Assessment of subjective status, as shown in Figure 3-1, may only be made by subjective reports. Subjective responses are a source of individual differences in workload capacity which frequently may be altered by appropriate perceptual structuring.

Individual differences in performance may arise from each of the three facets of the human operator discussed above (i.e., performance resources, physiological state, and subjective state). Because of the number and complexity of these facets, it may often appear impracticable to determine their relative impacts on performance differences. It appears clear, however, that the relative consistency of individual performance differences under fixed workload conditions (e.g., "workload 1" or "workload 2") may be directly determined through measures of reliability (i.e., $r(1,1)$ or $r(2,2)$). These metrics are of interest to the systems evaluator because they reflect on the importance of individual differences to system performance as well as the sensitivities of the performance measures to likely variations in workload (Sutcliffe, 1980). As with reliabilities, the consistency of performance measures across conditions (e.g., "workload 1" or "workload 2") may be determined by correlation of performances (viz., $r(1,2)$) and compared with the reliability measures just outlined. These cross condition correlations may also be pertinent to the system evaluator because of their reflection on the consistency of individual differences across workload conditions. Where a correlation is statistically consistent with the observed reliabilities, i.e.,

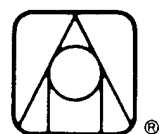
$$r(1,2) = \text{SQRT}[r(1,1)*r(2,2)] \quad (1)$$

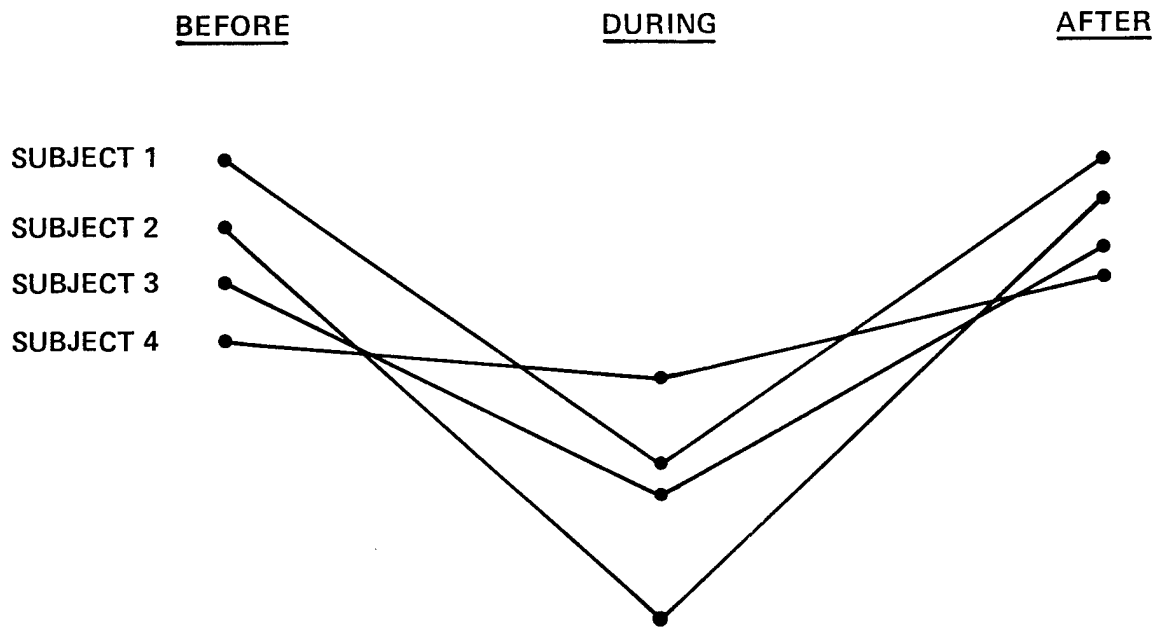


then it may be inferred that the individual differences across the two conditions are essentially isomorphic. This implies that individual differences are largely due to performance resources if one of the conditions (e.g., "workload 1") represents a selected condition of minimal loading.

When intercondition correlations are not consistent with the observed reliabilities (as illustrated in Figure 3-3), the indication is that the nature of the individual differences has changed. This could be due to changes in any of the three facets of the human operator discussed above (i.e., changes in the relative mix of performance resources used in performing, changes in physiological state, changes in subjective state, or a mixture of all three). Identification of the nature of such differential changes consequently requires assessments of each of the facets, so that their effects may be ascertained. The measures of reliability (correlations), as well as their comparison, are of interest to the systems evaluator.

It is clear that the identification of the nature of OWL individual differences (as resource, physiological, subjective, or some combination) ultimately rests on the demonstration of a relationship between these variables and performance. This requirement is important to keep in mind during the analysis of the extant literature as well as design and interpretation of future systems evaluation efforts.





	<u>BEFORE</u>	<u>DURING</u>	<u>AFTER</u>
BEFORE	1.0	-.1	.9
DURING		1.0	-.1
AFTER			1.0

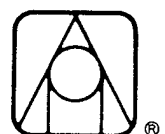
Figure 3-3. Individual Differences in Workload Response as Illustrated by Changes in Intercondition Correlation

4. REVIEW OF OWL MEASURES AND TECHNIQUES

4.1 TAXONOMY

The last thorough review of all available mental workload literature is believed to have been performed approximately six years ago by Wierwille and Williges (1978, 1980). In the 1978 document, they performed a survey and analysis of techniques in which 400 references were reviewed. They used a human operator task taxonomy (called Universal Operator Behaviors) that had been developed earlier by Berliner, Angell, and Shearer (1964) in which human activities in systems were separated into four categories: perceptual, mediational, communication, and motor processes. They also separated workload techniques and measures into four categories, namely subjective opinion, spare mental capacity, primary task, and physiological measures. Having performed these categorizations, they used them to determine which workload techniques had been applied to which operator behaviors, based on the research literature.

Other taxonomies of workload have been developed as well. These include subjective/objective, subjective/performance/physiological, and similar variants. In the years since the last known overall review of mental workload literature, research has continued and even accelerated. Two organizations have provided much of the impetus for the research: NASA Ames Research Center and the Air Force Aerospace Medical Research Laboratory (AMRL) at Wright-Patterson Air Force Base. NASA has done work both in house and through substantial external funding, whereas AMRL has performed most of its work in house, or with nearby contractor personnel. Of course, many other organizations have performed workload studies, both with and without external funding. As a result, many new results have been obtained.

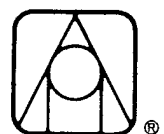


This section presents a specific workload taxonomy which, it is believed, is timely, flexible, and meaningful. The taxonomy is slightly different from previous ones, in that greater emphasis is placed on predictive techniques. Over the past several years, most research has involved empirical evaluation techniques rather than predictive techniques. As a result, most workload techniques are test-and-evaluation oriented, rather than preliminary-design oriented. However, it is quite clear that Army needs include both types of techniques. Front end, predictive techniques are needed for concept development and preliminary system design. Research in this area has not kept pace with need. In fact, very little recent research effort has gone into predictive techniques.

The workload taxonomy selected for this report breaks predictive techniques out as a separate category, called "analytical." The word analytical has been used to distinguish the category from other methods, which are termed "empirical."

Figure 4-1 shows the taxonomy selected. It becomes obvious in examining the taxonomy that empirical methods have received the bulk of emphasis in workload research. Numerous subcategories appear for empirical methods, whereas very few subcategories are identified for the analytical methods. The taxonomy is expandable and allows the addition of any number of subcategories. Many techniques that have been tried and found to be insensitive are not included, though they easily could be added later. The survey and analysis report done by Wierwille and Williges (1978) could be used to fill in those techniques. However, in this initial document, it was believed that representative, promising techniques should be highlighted in an effort to avoid unnecessary clutter and complication.

The remainder of this section of the report describes briefly each of the categories and subcategories listed in the taxonomy of Figure 4-1. References believed to be typical and to have created significant advances are cited. However, once again, only a very small portion of the available literature is cited, since it is estimated that at present, there are perhaps one thousand references on workload estimation in the literature.



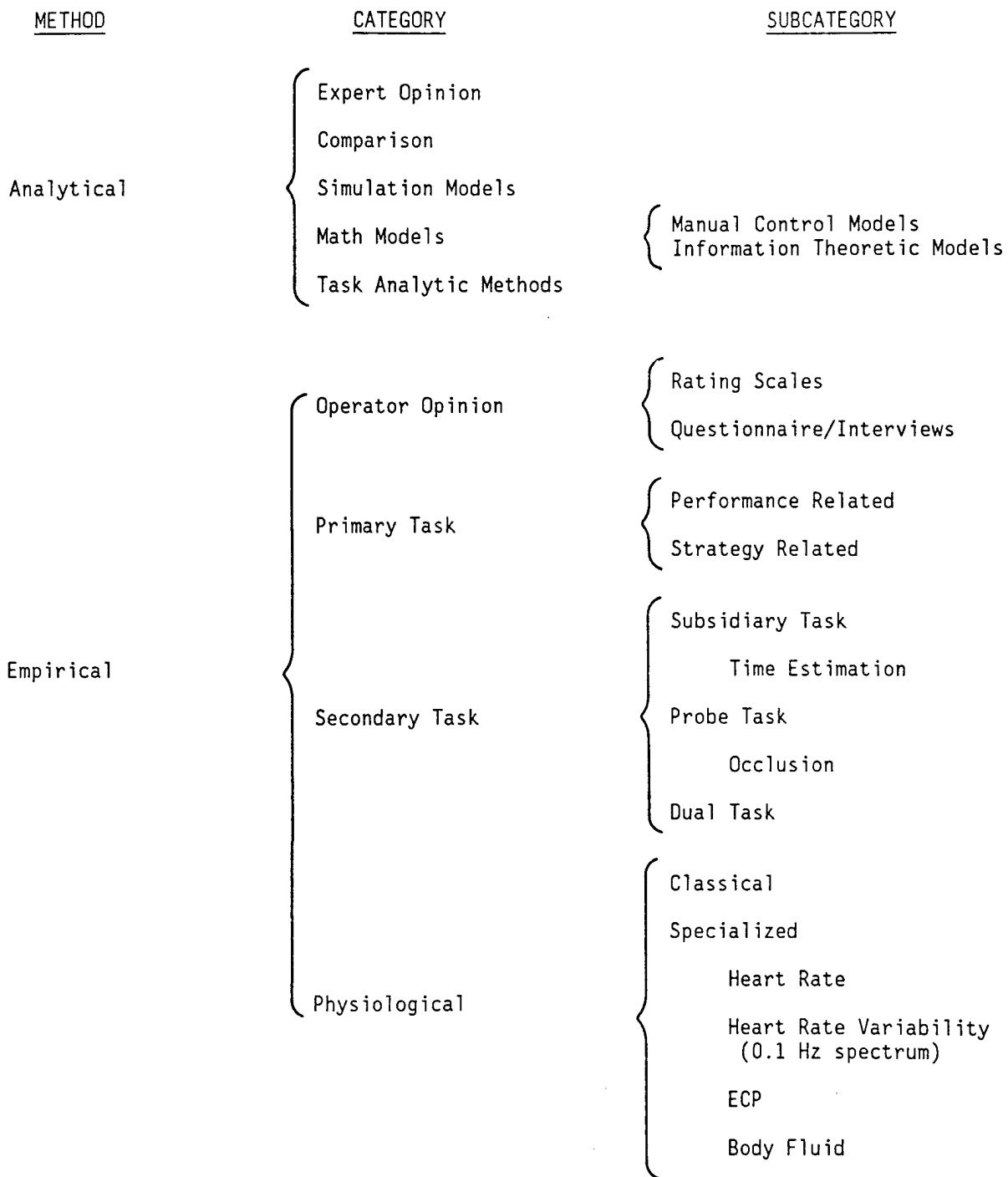
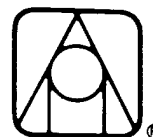


Figure 4-1. Expandable Taxonomy of Workload Assessment Techniques



4.2.2 Comparison

One of the best ways to evaluate aspects of any new system is to compare it to the closest available existing system. In terms of preliminary workload estimation, the same is true. If an operator workspace already exists that is similar to the one that must be designed, then that workspace provides an excellent means for estimating workload. If some of the empirical techniques described in Section 4.3 can be applied to the existing workspace to estimate its workload, then the estimates can be used as a basis for estimating the workload of the new interface. Adjustments can be made up or down as necessary to fit the new situation.

Of course, the comparison technique can only be used if there is an existing similar workspace. The specific application may differ substantially from the existing workspace, without having an appreciable effect on the preliminary estimation of workload. What is important is that the tasks be similar in character and that the requirements be similar in terms of such aspects as precision and response time. The less similar the tasks and aspects of the tasks, the greater will become the inaccuracy of the preliminary estimate of workload associated with the new workspace. However, as indicated earlier, the estimates can be revised up or down to account for some amount of dissimilarity.

Again, no reference or formal methods of transposing workload estimates from existing to preliminary new systems have been found. However, it is possible that references to use of comparison techniques for workload assessment were not identified because of the informal manner in which they are generally performed. We did learn of a case in which workload estimates for LHX scout missions, involving a one person crew, were obtained by comparison with a detailed workload analysis of scout missions in the OH-58D helicopter involving a two person crew (Shaffer et al., 1986 and Shaffer, personal communication); the method employed in this case appears to have been quite informal and subjective. Nevertheless, the comparison approach appears to be a valuable one and it could stand some formal development.



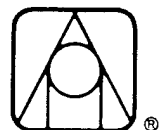
4.2.3 Simulation Models

Unlike the previous two categories of analytical techniques, there are numerous references to simulation models that can be used for workload estimation. Most of the major aircraft manufacturers have developed simulation models which in one way or another model human performance in systems. Other organizations have also developed simulation models which can be applied to systems other than aircraft. In other words, they are more widely applicable. Meister (1985) provides an excellent overview of simulation models, the great majority of which are now computerized.

Among the most widely used simulation models is the Siegel and Wolf (1969) model. The model simply identifies operator tasks and predicts their completion times and probabilities of success. It can be used in both a deterministic mode for design purposes and a stochastic mode for simulation purposes. The Siegel-Wolf model accounts for operator loading and therefore has ramifications for workload estimation in preliminary design. The effects of stress on performance can also be simulated, an important consideration for systems involved in combat and other forms of stress. According to Meister, the most recent versions of the Siegel-Wolf simulation model include task difficulty assessment, which would be of direct relevance to workload estimation.

SAINT is another simulation approach. It can be considered as an extension of the Siegel-Wolf model (Chubb, 1981). SAINT is actually a network programming language. It has wide application, including simulation of operator-machine systems. It has already been applied to several advanced systems including remotely piloted vehicles, advanced cockpits, and airborne warning and control systems. Being a network simulation, SAINT models the flow of operations in a system. It is unique in that it can allow branching based on outcomes of events.

Because SAINT is capable of gathering information on temporal aspects of performance, including task completion times, relationships between completion times, and task success, it can aid in the assessment of workload.



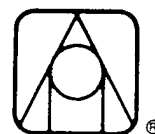
However, because of its generality, SAINT leaves a great deal up to the analyst in terms of defining what elements are to be assessed and how workload will be defined.

Perhaps the most sophisticated approach to simulation of operators in systems is HOS (Human Operator Simulator) (Harris et al., 1986; Lane et al., 1979). In using HOS, procedural descriptions are developed and used for simulation of the human operator, the system itself, and other systems having bearing on the operator and associated workspace. HOS, like the simulation programs of aircraft companies, can produce channel loading profiles as a function of time. In most cases, however, HOS is used to determine whether the operator can perform all of the tasks assigned in the time available. As such, it is similar to task analytical methods to be described in Section 4.2.5. However, HOS is a much more sophisticated method of simulation than are the task analytic methods, and is capable of providing much more microscopic structure to human operator tasks and task elements.

4.2.4 Math Models

In several cases, mathematical models are capable of serving as workload estimation tools for preliminary design. For a math modeling technique to serve appropriately, it must contain elements which reflect operator loading in some way. For example, if a math model is developed for human decision processes, then there must be some aspect or parameter of the model which changes as the load associated with the decision process changes. In other words, if the model does not adapt in some way to load, it cannot be used for workload estimation.

In this section, two quite different models will be presented. They have been selected because they meet the above stated criterion of reflecting loading and because they are widely applicable.



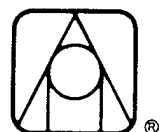
Manual control system models have been developed that make it possible to obtain a dynamic model of the human operator operating in a closed-loop system (Frost, 1972). The model of the operator can be made macroscopically accurate, such that the outputs of the model match rather closely those of the human operator. The model can be developed using transfer functions and closed-loop stability analysis methods, and is commonly known as the "describing function" of the human operator.

The basic components of the model include a delay of approximately 0.1 second, a lead-lag network for compensation, and additional high frequency roll off (in some circumstances) to account for limited bandwidth of the human operator.

Rules are available, based largely on classical control theory, that make it possible to determine the gain and the parameters of the lead-lag network. (The other parameters of the model are fixed.) Fundamentally, the idea is that the human operator sets gain and lead or lag so that certain open loop parameters are held constant, namely the unity-gain crossover frequency and the phase at crossover. Thus, the human operator model *adapts* to the system, maintaining closed-loop stability and performance.

Birmingham and Taylor (1954) had postulated earlier that human performance in a manual control system is optimum when the human operator need only exhibit gain and a short delay, that is, when compensation is unnecessary. As greater compensation becomes necessary by the operator, the system becomes more difficult for the operator to control. Thus, the amount of compensation the operator must provide is a measure of the difficulty the operator has in controlling the system.

Since the describing function allows a determination of the amount and nature of the compensation in the human operator model, it can be used to estimate workload. There is a direct connection between this compensation and the



workload that the operator will experience. As the amount of compensation increases, so does workload. Thus, model parameters directly reflect workload, as is necessary for the model to be helpful.

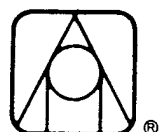
A second type of model is applicable to instrument panel monitoring. Senders (1964) developed a theory of attentional demand for situations in which the human operator monitors instruments on an instrument panel. The approach he used involved information theory. The equation for attentional demand which he developed was a summation of attentional demand of the individual instruments. To obtain the attentional demand of each instrument, he developed the concept of an ideal observer, one which would sample each instrument at the Nyquist sampling frequency. Sample duration was assumed to be proportional to the information generating rate of the instrument.

In Sender's formulation, the attentional demand associated with any given instrument is proportional to the bandwidth of the information multiplied by the logarithm of the allowable tolerance in reading the instrument. Thus, attentional demand is attainable directly from the model and meets the criterion described earlier for containing parameters which reflect workload.

In general, most mathematical models of human operators do not contain explicit parameters which reflect workload and therefore are of little value in obtaining preliminary estimates of workload. Modeling with the idea of workload in mind would seem to be a fruitful area for future research.

4.2.5 Task Analytic Methods

Task analytic methods are among the best available preliminary design techniques for estimating workload. The basic procedure is to first develop a mission profile which is typical of the way the system would be used. The profile then makes it possible to perform a task analysis for the operator of a given work station. This analysis, in turn, can be translated into activity profiles as a function of time.



The main objective is to uncover any situations in which the time required to complete any activities exceeds the time available. When this occurs, it can be assumed that the operator will be overloaded and will have difficulty in performing the required tasks.

Stone, Gulick, and Gabriel (1985) present an excellent overview of the techniques used by a major aircraft company to assess crew workload in the preliminary design stage. The techniques used are primarily task-analytic and they are applied in several different ways. Their method is computerized and has the following components:

Equipment interface workload -- This program is an overall workload metric. It sums each individual crew member's task times and relates them to the time available for each segment of a particular mission.

Body channel workload -- This program quantifies the overt physical actions taken by each crew member to determine level and amount of overlap. The action categories are verbal/aural, internal visual tasks, left-hand tasks, right-hand tasks, and foot tasks.

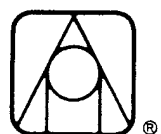
External vision availability -- This subprogram gives an indication of the amount of spare visual capacity available on the basis of running time. It is an indirect measure of the ability of the crew member to handle any emergency that might be superimposed on regular duties.

Any of these programs can also provide moment-to-moment task load on a relative basis. Consequently, "hot spots" can be detected and eliminated by redesign.

There is no question of the value of task analytic methods. They have evolved over many years and have been shown to be very helpful in assessing workload in preliminary design. They should certainly receive emphasis in the Army's development of an OWL battery.

4.3 EMPIRICAL TECHNIQUES

As indicated earlier, the great majority of workload research performed in the past two decades has been on empirical evaluation. As a result, the research literature has numerous citations pertaining to empirical methods, many



of which have reached the stage where they are now relatively well proven. In this section, some of the more important findings are described briefly.

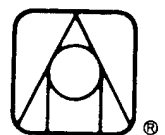
4.3.1 Operator Opinion Measures

Operator opinion is often used as a workload evaluation measure. Sometimes, the measure is used in conjunction with other measures, as a means of gaining insight and perspective on workload problems. There are many possible techniques for gathering opinion, including psychometrically derived rating scales, dichotomous and multiple choice responses, structured and unstructured questionnaires, and structured and unstructured interviews.

There are two general approaches contained within the above techniques. One of these is fundamentally to obtain "ratings" in which an attempt is made to quantify on an ordinal or interval basis, one or more workload-related attributes of a system. In general, this approach involves ranking of alternatives, selection of a point on a continuum, or assignment of a numerical value. The second approach is the questionnaire/interview approach. In this case, the information obtained is qualitative. Interviews and questionnaires are not based on scaling considerations, but nevertheless can be highly valuable. They allow the operator to provide impressions of workload-related elements and can be used to bring out subtleties that might be very difficult to capture using rating techniques.

There is little question that subjective impression is an important aspect of workload. If operators did not experience workload, then opinion methods would be of little value. Because workload is largely an experience, opinion methods are not only desirable as a method of assessing workload, they are almost essential.

For *rating* scales, there have been two main stems of work. The first of these is what might be termed the "engineering" approach to scale design. The best example is the Cooper-Harper (1969) scale. This is a 10-point decision

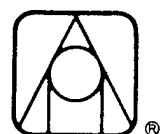


tree scale that was derived through experiment and intuition to allow the rating of handling qualities of aircraft. The scale does contain some references to workload related elements and has been used in some cases as a measure of workload.

The engineering scale approach has continued to the present day. Examples include the so-called Bedford scale (Roscoe and Grieve, 1986) which is similar to the Cooper-Harper, but with emphasis on "spare capacity," and the Modified Cooper-Harper (MCH) scale (Wierwille and Casali, 1983), which emphasizes workload and errors. Both scales have been tested effectively and have been found sensitive to task loading along a variety of dimensions. Also the MCH scale has been compared against variants of itself including 15-point scales, computerized scales, and scales without decision trees (Skipper, Rieger, and Wierwille, 1986). It was found that the standard MCH scale had greater consistency than any of the variants. It has been established that an engineering intuitive approach to workload rating scales is a valuable one.

A wide variety of psychologically derived scales have been devised. These range from a relatively intuitive approach to a psychometrically derived approach. In fact, general references are readily available on rating scale construction and other scaling considerations (Edwards, 1957; Nunnally, 1967). These are "attitude scale" construction references, and do not deal specifically with workload estimation. Nevertheless, the techniques presented do have bearing on scale construction for workload estimation.

Most recently, two new scales have evolved, based largely on psychometric techniques. The first of these is a set of bipolar scales known, appropriately, as the NASA Bipolar ratings (Hart, Battiste, and Lester, 1984). The second is called SWAT (Subjective Workload Assessment Technique) (Reid, Shingledecker, and Eggemeier, 1981). Both scales account to some extent for both the multidimensional nature of workload and individual differences.



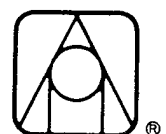
The NASA-Bipolar technique involves two steps: workload parameter evaluation and actual data gathering. The scale has nine aspects which it assesses: time pressure, task difficulty, performance, mental effort, physical effort, stress, fatigue, frustration, and activity type. In the workload parameters evaluation phase, the subject compares the aspects in terms of importance, and a weighting is then computed. Subsequently, this weighting is used in the data gathering phase to weight the subject's aspect ratings, thereby achieving an individualized rating of workload for the task at hand.

SWAT is also a two-step process. The scale is made up of three dimensions, each having load levels of low, medium, and high. The dimensions are time load, mental effort, and psychological stress load. Because of the 3 by 3 by 3 nature of the SWAT, there are 27 possible combinations that can be obtained. In the scale development phase, the subject ranks the 27 possibilities in an ordering process. Then, in the data gathering phase the subject rates each single dimension as low, medium, or high.

SWAT is unique in that it takes advantage of conjoint scaling techniques to produce what is claimed to be a transformation of the three dimensional ratings into a single *interval* metric.

The MCH, the NASA Bipolar, and SWAT rating methods have all been thoroughly tested. They are quite sensitive to a wide range of operator loading dimensions, and therefore, they provide good metrics of the subjective impression of imposed workload.

In terms of the questionnaire approach to evaluation of workload, the major aircraft companies have developed questionnaires for aircrew workload estimation. Typically, these questionnaires contain specific questions to which the aircrew member is to write an answer. In most cases, the questionnaires are administered after completion of a flight phase. Examples of these appear in



McLucas, Drinkwater, and Leaf (1981), a report on the results of crew complement hearings, and in Ruggiero and Fadden (1982).

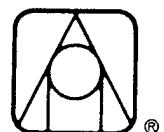
These questionnaires rely heavily on the experience and technical knowledge of the aircrew member. On the other hand, the questionnaires provide substantial freedom in terms of answering. Consequently, they serve the extremely useful, perhaps even unique, purpose of allowing the uncovering of workload problems that could not be detected using ratings or other methods of workload.

In general, operator opinion data are among the easiest to obtain. An experimenter need only administer the necessary forms and follow instructions (or procedures) carefully. Part of the popularity of opinion data is a result of its practicality and ease of use.

4.3.2 Primary Task Measures

As the mental workload experienced by a human operator increases, it is likely that the operator's performance will change. While moderate loading may not cause large changes in performance because of operator adaptation, loading of sufficient magnitude will certainly cause changes, usually in the direction of degradation. When such changes do occur, they can be measured and considered as indicative of heavy workload.

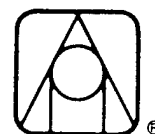
Primary task measures have not been particularly popular during the last decade as indicants of workload, because, as Cooper and Harper (1969) put it, "In a specific task, he [the pilot] is capable of attaining essentially the same performance for a wide range of vehicle characteristics, at the expense of significant reductions in his capacity to assume other duties." In other words, performance measures as indicants of workload, are "latent" indicants. Loads may have to be excessive to cause measureable changes. Nevertheless, as such they might be quite valuable. If operators in a given system seem to be making errors at a high rate, or if other measures of performance are suffering, high operator loading may be present.



Performance measures are by no means the only forms of primary task measures of workload. Other measures, can also be obtained which might exhibit changes at much lower levels of loading. In general, when an operator's task becomes more difficult, that operator musters more resources to maintain performance at what is perceived to be an acceptable level. This increase in effort must manifest itself as some change in the operator's output; in other words, the operator's strategy shifts to meet the demands of the task. If a measure can be devised which tracks this shift in strategy, then that measure will reflect workload.

To contrast performance and strategy, consider a situation in which an operator is performing a tracking task. The operator strives to hold error within some tolerance. If the order of the tracking task is increased, the dynamics become more difficult to control. The operator then must apply more anticipation or lead to control the system. Such lead shows up as sharper rise times in the operator's responses and as phase lead in the operator's describing function. If a measure is devised which assesses this lead, the measure will reflect operator loading, even though operator-machine performance may remain within tolerance. Thus, primary task measures based on strategy or similar concepts may be very reliable measures of workload, contrary to what some researchers believe. However, if only *performance* measures are analyzed, the likelihood is high that they will not reflect changes other than overload or near overload.

Primary task measures are among the most underrated of workload estimation measures. Wierwille, Casali, Connor, and Rahimi (1985) ran four different experiments, one for each of the major operator behaviors presented in the Universal Operator Behaviors taxonomy (Berliner, et al., 1964). They found in each experiment that at least one primary task measure which tapped strategy in each experiment was as sensitive as any other measure taken, including rating scales. Consequently, it is clear that sensitive primary task measures based on strategy can usually be found.



Primary task measures are often quite easy to obtain when an operator performs in a simulation. Usually there is a means by which the appropriate signal can be tapped and analyzed. Often there is even computational power available to allow measures to be computed. Where actual equipment is involved, however, (as opposed to simulation), it becomes more difficult to gain access to the necessary signal or parameter. Usually sensors must be added and signals from them conditioned. Thereafter, computational power must be made available to obtain measures. However, with modern microcomputers and a great variety of interface cards available, measures computation has become straightforward in almost every situation. Therefore, primary task measures, particularly those based on strategy changes, appear very promising.

4.3.3 Secondary Task Measures

The secondary task approach to workload measurement involves the evaluation of spare (residual or reserve) capacity. In classical terms, the operator is assumed to function as a single channel processor (Knowles, 1963; Rolfe, 1973). Spare capacity is then assumed to be the difference between total workload capacity and that required by the particular (primary) task at hand. In theory then, it should be possible to introduce an additional or secondary task which the operator performs whenever spare capacity permits, thereby bringing the operator up to full capacity. Performance on the secondary task should improve as primary task demands decrease, and should degrade as primary task demands increase.

Human operators are, in fact, not single channel processors and they also do not have a fixed upper limit to capacity. In modern concepts, operators in systems are viewed as having resource pools that can be drawn upon in different ways (Kahneman, 1973). For example, communications can be carried on in most cases while performing status monitoring of a system. On the other hand, communications might be very difficult to carry on while attempting to perform even simple mental arithmetic operations. As soon as cognitive load increases, communications must be time-shared.



This shows why secondary tasks may or may not be sensitive to various kinds of operator loading. An operator could conceivably be performing a primary task drawing on one subset or domain of resources while performing a secondary task that draws on a mutually exclusive second subset or domain. Under such circumstances, any change in difficulty of the primary task may cause no corresponding change in the secondary task. In other words, the secondary task is not sensitive to the changes in loading of the primary task.

On the other hand, if a secondary task requires a domain of resources that substantially overlaps the domain needed to perform the primary task, then the secondary task may become sensitive to primary task loading, but may also cause intrusion. Intrusion can be defined as an undesirable change in primary-task performance, resulting from the introduction of a workload estimation technique or apparatus. The trick then is to select a secondary task that is sensitive for the specific primary task to be measured while not intruding to any appreciable extent.

A great deal of dual task research has been done in the past five years which is aimed at determining what relationships exist between tasks of various kinds. (See, for example, Tsang and Wickens, 1984.) This research appears somewhat esoteric, but in fact has bearing on the selection of a secondary task to be used with a specific primary task. Unfortunately, this research has not reached the stage where it can be used to make specific recommendations. Part of the reason for this is the two-dimensional nature of the tasks. In other words, the number of possible combinations in dual tasks is staggering.

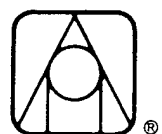
In recent years, numerous secondary tasks have been tried in a variety of experimental conditions in an effort to determine their ability to measure workload without intrusion. Among the techniques examined have been digit shadowing, Sternberg paradigm (memory scanning), mental arithmetic, Michon tapping, and time estimation. Only two of these have proved to have any general worth as indicants of workload: memory scanning and time estimation. Furthermore,



memory scanning seems to suffer from lack of generality. Certain investigators have found it to be sensitive while others have not. As a result, it appears that it cannot be relied upon for general applications. On the other hand, time estimation appears sensitive, robust, and relatively non-intrusive. Wierwille, et al., (1985) found the standard deviation of time estimation to be sensitive to operator loading in four different experiments. Hart (1975) had also found time estimation to work well. Hart indicates that the probable reason for the sensitivity of time estimation to load is that cognitive processing interferes with the ability to estimate time accurately. As cognitive load increases, the operator must assess time by the number of other events that have occurred. In so doing, the variability of the estimates increases. One negative note on time estimation as an indicator of workload is that, although time estimates may exhibit good reliability within a session, they have been shown to be unstable across sessions. (McCauley et al., 1980; Bittner et al., 1986).

The research on secondary tasks during the last ten years has shown that secondary tasks generally cannot be relied on for workload evaluation, with one exception -- time estimation. This technique seems to work well in a variety of applications and it is easy to implement. In its simplest form, operators can be instructed to press a button or speak a word ("now" for example), say, once every ten seconds while performing a given primary task. The intervals can be timed and the measure computed from them. A better procedure is to provide a stimulus which marks the beginning of the interval and have the subject respond at the end of the interval. This procedure tends to reduce confusion regarding the beginning and ending of intervals.

Thus far in this section, only *subsidiary* tasks have been discussed. In a subsidiary task, the operator performs the primary task and only performs the secondary task when the demands of the primary task permit. In other words, the secondary task has a lower priority than the primary task. There is, however, another class of secondary tasks which could be called "probe tasks." In this case, the operator is instructed to perform the secondary task regardless of primary task demands. The difficulty of the secondary task is

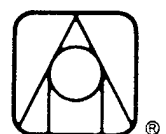


then increased until primary task performance begins to suffer. The level of difficulty of the secondary task, at this point, is then a measure of the spare mental capacity of the operator, and therefore, indirectly, a measure of workload of the primary task (Knowles, 1963).

Probe tasks have limitations because by definition they must intrude to some extent on the primary task. In the operation of ordinance, vehicles, and certain systems, such intrusion for purposes of measurement could not be permitted for safety reasons.

One probe-type secondary task is that of occlusion, which involves forced visual sampling. The usual method for accomplishing this is to block the operator's visual input from the system. Blocking is accomplished in one of two ways. One way is to have the operator wear a helmet or hat fitted with an opaque visor. The visor is electrically or pneumatically controlled. It rapidly flaps down into the operator's field of view, remains there for a specified time, and then flips up again. The other way of blocking the visual input is to blank the displays of the system directly. This method is usually limited to projection, CRT, electroluminescent, LED, and similar electronic displays in which electronic control over blanking is possible. As the time between visual samples increases, degradations in primary task performance begin to occur. Thus, time between samples serves as an indirect metric of primary task load (Senders, Kristofferson, Levison, Dietrich, and Ward, 1967; Farber and Gallager, 1972).

This method of evaluation requires special equipment, but such equipment can generally be fabricated without difficulty. The method itself is only applicable to visual processes, however, and even then suffers from the limitations described above.



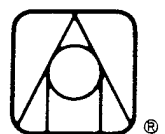
Mention has already been made of dual tasks as an approach to workload. In particular, it has been indicated that they may be helpful in selecting a subsidiary task for workload. However, dual tasks may also serve as a method for estimation of workload directly. If, for example, a system operator's duties can be dichotomized in some way into two sets of subtasks, then measures of each subtask can be developed. Accordingly, there may be shifts in either or both of these measures as a function of increased load. There has been little application of dual tasks in this way to assess workload, but the concept is mentioned here for the sake of completeness.

4.3.4 Physiological Measures

Physiological measures of workload have received a great deal of attention during the last two decades. Almost every researcher in the field has dealt with these measures in hopes of finding a "reliable, objective, sensitive" measure of operator loading. Underlying physiological measures of workload is the idea that as operator loading increases, involuntary changes take place in the physiological processes of the human body (body chemistry, nervous system activity, respiration, and circulation). Thus, workload may be assessed by measuring the appropriate physiological variables and processing time.

The reason for the direct connection between workload and physiological measures has often been assumed to be "arousal." Arousal may be considered to be a state of preparedness of the body or level of activation of the human organism.

Closely associated with physiological measures of workload is the concept of stress. Essentially, high workload is assumed to induce stress which in turn causes changes in physiological measures. Consequently, it is assumed that some physiological measures may be directly affected by workload, while others are affected indirectly through an increase in stress.

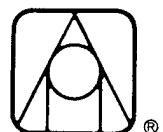


Early work on physiological measures of workload generally followed from the medical equipment availability for measuring heart rate, blood pressure, brainwave activity, skin conductivity, and various chemical compound levels in the blood and urine. In most cases, these early studies indicated that physiological measures generally are not sensitive to workload, but may be sensitive to several other forms of stimuli, such as physical exertion, environmental conditions, and diet. Wierwille (1979) in a summary article on physiological measures of workload stated, "While a wide variety of physiological measures have been developed for assessment of workload, few if any of them are at present proven to the extent that they can be widely applied"

Since that time, work has continued which does show promise for specific situations. Roscoe and Grieve (1986) continue to advocate (as well as demonstrate) that pilot workload is capable of being assessed by heart rate. Wierwille, et al., (1985) also conclude that heart rate will increase reliably under conditions of perceived danger, which can accompany increased workload in systems.

Very recently, Morey et al. (1986) have shown that the 0.1 Hz component of the spectrum of the interbeat interval of heart rate is indeed sensitive to operator loading. This work is based on the earlier work of Mulder (1973) and Sayers (1973). Until Morey et al. performed their studies, heart rate variability was controversial because many investigators had negative findings. It now appears that the spectral component about the 0.1 Hz point is sensitive to cognitive load and may therefore provide a reliable measure of workload.

Another interesting and relatively recent development has been the connection of certain types of cognitive processing with portions of the ERP (Evoked Response Potential) or ECP (Evoked Cortical Potential). (See, for example, Kramer, Wickens, and Donchin, 1983). The fundamental concept is that the P300 portion of the ERP waveform and cognitive processing are related. For example, if an operator is viewing a stimulus while cognitively processing it,



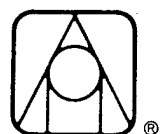
the amplitude of the P300 is different from the case in which the operator is observing but not processing the information. Thus, the P300 is, to some extent, a measure of cognitive load associated with a stimulus, and perhaps attention. Numerous papers have been written on the P300 component and the kinds of task elements and cognitive processing to which it is sensitive. The P300 of the ERP represents a specialized form of workload assessment that could possibly be used in particular kinds of workload estimation. The method applies only to those situations in which there is a repetitive, discrete stimulus to which the ERP can be synchronized. Thus, as examples, a display must flash new information or an auditory tone must be initiated repeatedly before sufficient data can be gathered.

Finally, it should be mentioned that body fluid analysis can serve as an effective workload indicator in certain circumstances. Changes are known to take place in the body's metabolic balance when undergoing stress or becoming fatigued. Since prolonged workload can induce stress or fatigue, it becomes possible to relate workload to changes in the chemistry of body fluids.

Hale, et al. (1974) have developed a stress battery which includes measurement of seven compounds in the urine or in the parotid fluid:

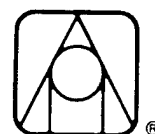
- norepinephrine -- nervous system activity
- epinephrine -- adrenomedullary activity
- 17-OHCS -- adrenocortical activity
- urea -- protein catabolism
- sodium -- mineral metabolism
- potassium -- mineral metabolism
- sodium to potassium ratio -- metabolic balance

Changes in these compounds reflect the levels of activities listed after each compound. The main difference between the body fluid approach to physiological measures of workload and virtually all other measures is that it reflects the



long term effects of workload, say over several hours. All of the other techniques assess workload over intervals of perhaps a few minutes. For Army applications, the body fluid approach could become quite valuable because of this unique "long term" property.

In general then, certain specific physiological measures do reflect specific types of operator loading. As such they can be useful in special circumstances. Equipment needed for each identified physiological measure is unique. Thus, only one measure can be obtained with a specific equipment complement. Nevertheless, if a measure is needed for the specific type of loading that a given physiological measure can provide, the specialized equipment may be worth the cost.



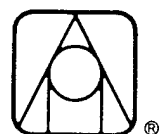
5. OWL MANAGEMENT INFORMATION SYSTEM (OWLMIS)

An OWL Management Information System (OWLMIS) is under development to provide for the control and analysis of the mass of documents and other resources comprising the OWL Scientific Base (ca., 1000 items). Currently under implementation, the system is composed of three components: (1) the OWL Information Data Base (OWLIDB); (2) the OWL Information Analysis System (OWLIAS); and (3) the OWL Library (OWLLIB). The OWLMIS has been designed to provide for both efficient transfer and analyses of the OWL Scientific Base.

5.1 OWL INFORMATION DATA BASE (OWLIDB)

The OWLIDB is structured to be an evolutionary system for efficient analysis of the OWL Scientific Base. Using an efficient (1000 byte, 13 field) citation format, bibliographic information is encoded which enables a user to locate documents or resources by combinations of: author; key term (in title) context; producing organization; publisher; and publication date (or range of dates). Additionally, within fields expected to occupy well less than a maximum of 1000 bytes, each citation will be annotated on dimensions required for scientific base analyses. These presently include:

- Measures. -- In the terms of the taxonomy presented earlier (Figure 4-1), which set(s) and subset(s) of OWL measures are discussed?
- Individual Differences -- Does the document address the issue of subject differences, stability, or reliability-efficiency (Bittner et al., 1986) directly or does it contain relevant data? (Although not well recognized, reliability and stability estimates frequently may be derived from summary statistics for repeated measures data, e.g., Fs and error terms.)



- Sensitivity -- Does the document contain information on the comparative sensitivity of measures of workload? (Care must be taken to consider sensitivity-efficiency, the parallel of reliability-efficiency.)
- Validity -- Does the document address the validation of measures of OWL?
- Report Nature -- Is the document theoretical, a review, laboratory research, a systems application, or a DoD policy document (e.g., MIL-STD, MIL-SPEC, etc.)?

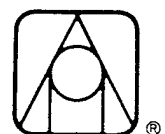
These annotations provide for the ready identification of reports which deal with issues concerning the systems applications of measures of OWL. The OWLIDB is designed to provide for efficient (i.e., less than 2000 bytes per citation) bibliographic and salient systems applications analyses.

5.2 OWL INFORMATION ANALYSIS SYSTEM (OWLIAS)

The OWLIAS is composed of the software and hardware components used to build and exercise the OWLIDB. Initially, the building of bibliographic and annotation information is being accomplished using relational data base software (dBASE III) on an IBM PC compatible (MS DOS) system. Anticipating the use of the OWLIDB for preparation of publications in a variety of reference formats, it is also noteworthy that our dBASE III structure is compatible with commercial bibliographic (reference) software which are being explored for purposes of augmentation of our extant system. The hardware environment was selected to be broadly compatible with government standard microcomputer systems. The OWLIDB, as currently configured, would require less than 2MB of total storage and could be comfortably installed on a 5MB hard disk. The OWLIAS and OWLIDB components have been designed to provide for easy transfer of the OWLMIS to government and other facilities.

5.3 OWL LIBRARY (OWLLIB)

The OWLLIB is being assembled from formal and informal sources. These sources are described in the following along with the status of the OWLLIB.



5.3.1 Formal Sources

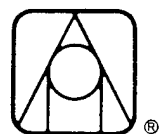
Computer-based searches of the literature have been initiated. Such searches involve a close collaboration of project personnel with a professional information scientist. Bases being searched include:

- Books containing appropriate key search terms in the titles or as descriptors, available from the publisher or in the Library of Congress catalog;
- Unpublished documents in the Library of Congress;
- Books and other documents, available on loan or through a copy service from the British Library;
- Dissertations;
- Reports on government funded research;
- Publications available through DTIC and NTIS;
- Conference papers;
- Conference proceedings;
- Journal articles;
- Monographs and other unpublished reports.

Boolean searches will be accommodated over a number of fields (e.g., corporate author A, between years B and C, but not topic D), including terms imbedded (i.e., a global search for a term that might be part of an entry).

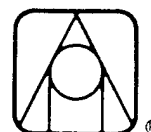
5.3.2 Informal Sources

A primary source of information is the "informal university," based on personal contact with people engaged in the investigation of operator workload. Leads on research, planned or in progress, have been and will be solicited and the principals contacted in person. Information from these sources are being collected and maintained in the OWL library.



5.3.3 Current Status

The OWLLIB is currently being assembled to support the principal reference needs of project personnel. It now contains over 500 citations and 100 documents, with about 1,000 more raw citations awaiting processing, possible classification, and entry. It is anticipated that by the time Task 2 is completed, there will be approximately 300-500 physical documents in the library.



6. SUMMARY AND PLANS

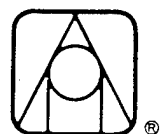
6.1 SUMMARY

The preceding sections have presented the principal elements of the foundation of the OWL program (i.e., Task 1). OWL was defined as the relative limitation in the human's ability to do work, rather than as some single- or multi-dimensional quantity which might be directly measurable. An OWL performance model was described in order to relate OWL to all other key aspects of human and system performance and so provide a conceptual tool to support subsequent efforts in the OWL program. A review was then presented of currently available tools and measures for the assessment of OWL along with a comprehensive taxonomy broken into the major categories of analytical methods and empirical methods. Finally, strategies and tools for the management of documents and information pertinent to the OWL program were described.

Four additional tasks are planned to complete the OWL program. These tasks build on the results of Task 1 to meet the ultimate objective of providing the Army with validated methods and handbooks for OWL assessment through all phases of combat system development. Fairly detailed plans for Task 2 are presented here, followed by brief descriptions of plans for the subsequent tasks.

6.2 TASK 2 PLANS

With the completion of Task 1, our immediate attention now turns to Task 2 and our approach rapidly shifts gears from the academic to the more pragmatic. Task 2 initiates the concentrated effort to marry the philosophical world of workload constructs to the U.S. Army's world of reality. That reality



contains a critical requirement for practical predictors and metrics of the workload concept which we have defined earlier in this Task Report.

Task 2 comprises the identification of formal Army requirements regarding workload, the assessment of user needs, the selection of specific systems to serve as "test cases" for evaluation and validation of selected workload assessment techniques, and the initial outlines of the OWL handbooks.

6.2.1 Formal Army Requirements

Table 6-1 repeats the list of relevant documentation initially identified in our proposal. Many of these documents have already been obtained and their review has been initiated during Task 1. In addition, other pertinent titles, information, and documentation have subsequently been obtained from the COTR and have proven very useful. The "1986-1987 Army Command and Management: Theory and Practice" document has provided a complete and timely look at Army systems management practices. Chapter 17, in particular, discusses many of the points we have previously raised, and provides insight into the the relevant portions of many of the documents we originally identified as germane in Table 6-1. More importantly, however, it identifies additional reference material which will enrich our reviews and conclusions from Task 2. A major section of the Task 2 Interim Report shall be a condensation and amalgamation of this comprehensive literature review. Six to ten of the most meaningful documents shall be selected and discussed as to why they are most critical, and what their potential is to impact down-stream system design and user acceptance.

6.2.2 Assessment of User Needs

The assessment of user needs shall be conducted as we originally proposed; i.e., interviews shall take the inadequacy and lack of versatility of current workload concepts as a given. Proceeding from that point, we will be more easily able to focus on realizable solutions. A final list of organizations



Table 6-1. Relevant Documents

DEPARTMENT OF DEFENSE INSTRUCTIONS:

DoD Directive 5000.1 Major Systems Acquisition
DoD Directive 5000.2 Major Systems Acquisition Process
DoD Directive 5000.3 Test and Evaluation

MILITARY STANDARDS:

MIL-STD-1472C Human Engineering Design Criteria for Military
Systems, Equipment, and Facilities

MILITARY SPECIFICATIONS:

MIL-H-46855B Human Engineering Requirements for Military Systems,
Equipment, and Facilities

MILITARY HANDBOOKS:

MIL-HDBK-759 Human Factors Engineering Design for Army Materiel

ARMY REGULATIONS:

AR 10-41 Organization and Functions, U.S. Army Training and Doctrine Command
AR 15-14 Systems Acquisition Review Council Procedures
AR 70-1 Army Research, Development, and Acquisition
AR 70-10 Test and Evaluation During Development and Acquisition
AR 71-3 User Testing
AR 602-2 Human Factors Engineering Program
AR 1000-1 Basic Policies for Systems Acquisition

Table 6-1. Relevant Documents (continued)

TRAINING AND DOCTRINE COMMAND PUBLICATIONS:

TRADOC-R 600-4 Guide for Obtaining and Analyzing Human Performance Data
in a Materiel Development Project

TRADOC-PAM 351-4(T) Job and Task Analysis Handbook

DATA ITEM DESCRIPTIONS:

DI-H-7054 Human Engineering Systems Analysis Report

DI-H-7055 Critical Task Analysis Report

DI-H-7056 Human Engineering Design Approach Document -- Operator

DI-H-7057 Human Engineering Design Approach Document -- Maintainer

AERONAUTICAL DESIGN STANDARDS:

ADS 30 Human Engineering Requirements for Measurement of Operator Workload
(DRAFT)



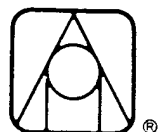
to be interviewed shall be jointly prepared with the COTR, and a schedule arranged. Subjects to be addressed during the interviews shall include:

- System acquisition cycle
- Who is the "designer"?
- Who is the "user"?
- Major overall OWL needs
- The adequacy of the workload concepts contained within this report
- Projected additional user needs likely to emerge and how to address those projected needs
- Workload problems in specified Army systems or types of systems
- The "ideal" workload measure
- How to transform OWL assessments into system design inputs
- Review of preliminary handbook outlines
- Possible systems to serve as our test cases

Where possible, both group and individual interviews shall be conducted. Our experience is that the group interviews often provide a synergy which produces concepts or conclusions unavoidably overlooked or vaguely formulated by individual contributors. Follow-up interviews shall be scheduled with all organizations and individuals approximately 12 months subsequent to the initial contact.

6.2.3 Outline of Handbooks

As we have repeatedly stressed throughout our proposal and initial effort on this project, UTILITY is the primary goal of the handbook development. The excellence of the material in the handbooks and the mathematical rigor of the measurement techniques and algorithms proposed are useless if the handbooks are not a tool which the designer or user can relate to, employ with a minimum



of confusion, and trust to produce a valid indicator of system workload. Handbooks shall be outlined within Task 2 which tell the Army user:

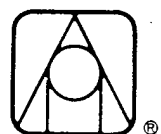
- Where he is in the system development cycle;
- What techniques are available to him;
- How to judge the usefulness of existing workload data;
- How to estimate the quality, quantity, and format of required data;
- The best course of action in light of the prevailing situation.

The outlines themselves shall comprise complete table of contents, a narrative describing the intent of the section, probable length, approach taken with each particular subject, recommendations for additional subjects or issues whose inclusion is tentative prior to the interviews, and preliminary ideas for graphical composition, formatting, and charts or nomograms. Additionally, some consideration will be given to the possibility of presenting the handbook information not only in hard copy, but alternative media, including optical disk or computer based interactive.

6.2.4 Selection of Specific Systems

The last item of business to be completed within Task 2 will be the selection of representative Army systems. Our recommendations shall be formulated and developed as an iterative exercise in parallel with the Army; i.e., there will be no surprises presented within the Task 2 Interim Report. Our recommended systems shall be selected upon the following criteria:

- What systems does the Army feel are prime candidates?
- What systems does our combined team have the richest experience base with?
- What systems possess the best data base of real or postulated workload measures?
- What systems can be realistically impacted (improved) by our experimentation?

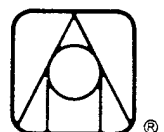


Task 2, with the exception of the follow-up interviews, is scheduled to be completed six months after date of contract. Its scope and execution is just as ambitious as those experienced in Task 1; the interviews, the travel, and the compilation and consolidation of subjective data will be extremely challenging. The rewards, however, are apparent. This will be our initial contact with the real world user community. We expect to gain a rapid education concerning the practical impediments to predicting and measuring workload. In fact, we will gain far more from the Army's teaching during this task than the Army will from us; we will repay this debt during Tasks 3, 4 and 5 when we demonstrate that we have listened objectively and responded with the development of a practical, rather than merely an elegant solution.

6.3 PLANS FOR SUBSEQUENT TASKS

Task 3 will be concerned with evaluation of measures of OWL and developing a validation plan for the OWL methodology and for the component OWL measures. We will evaluate the utility of the OWL measures identified in Tasks 1 and 2 relative to the system design and evaluation needs of TRADOC, AMC, and OTEA. Validation plans will be prepared for all recommended OWL measures focusing particularly on the prototype systems recommended and approved in Task 2. Plans for OWL analyses of the prototype systems will also be produced.

In Task 4, the plans generated in Task 3 will be implemented. We will conduct studies to validate the OWL measures and methodology and we will perform OWL analyses of the prototype systems. It is anticipated that some validation studies will be conducted as experimental efforts separate from the prototype system analyses, though the OWL assessments of the prototype systems will also provide important validation data for all aspects of the OWL methodology. In addition to demonstrating and validating the OWL methodology, it is intended that the analyses of the prototype systems will provide significant direct benefits for the development of those systems.

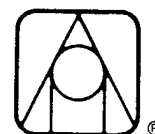


Task 5 will address the production of the primary documentation of the OWL program. We will develop the TRADOC pamphlet, Workload Estimation Handbook, and Workload Evaluation Handbook according to the outlines generated in Task 3. A Post-Contract Survey form will be prepared to provide an efficient vehicle with which TRADOC, AMC, and OTEA can assess the degree to which the other OWL documents have met their needs. In addition, a technical report detailing the scientific basis for the information contained in the pamphlet, handbooks, and form and discussing further Army research in the area of controlling operator workload will be prepared.



REFERENCES

- Berlinger, C., Angell, D., and Shearer, D. J. (1964). Behaviors, measures, and instruments for performance evaluation in simulated environments. Albuquerque, NM: Paper presented at the Symposium and Workshop on the Quantification of Human Performance.
- Birmingham, H. P., and Taylor, F. V. (1954). A human engineering approach to the design of man-operated continuous control systems. Washington, D.C.: Naval Research Laboratories, Report No. NRL-4333 (April).
- Bittner, A., Carter, R. C., Kennedy, R. S., Harbeson, M. M., and Krause, M. (1986). Performance evaluation tests for environmental research (PETER): evaluation of 114 measures. Perceptual and Motor Skills, 63, 683-708.
- Casper, P. A., Shively, R. J., and Hart, S. G. (1986). Workload consultant: A microprocessor-based system for selecting workload assessment procedures. Proceedings of IEEE/SMC International Conference. Atlanta, GA: IEEE/SHC.
- Chubb, G. P. (1981). SAINT, a digital simulation language for the study of manned systems. In Manned System Design: Methods, Equipment, and Applications, Moraal, J. and Kraiss, K. (eds.). New York: Plenum Press. 153-179.
- Cooper, G. E. and Harper, R. P., Jr. (1969). The use of pilot rating in the evaluation of aircraft handling qualities. Moffett Field, CA: NASA-AMES Research Center, Report No. NASA TN-D-5133. (April).
- Edwards, A. L. (1957). Techniques for attitude scale construction. New York: Appleton-Century-Crofts.
- Farber, E. and Gallagher, V. (1972). Attentional demands as a measure of the influence of visibility conditions on driving task difficulty. Highway Research Record, No. 414, 1-5.
- Frost, G. (1972). Man-machine dynamics. In Human Engineering Guide to Equipment Design, Revised Edition, Van Cott, H. P. and Kinkade, R. G., (eds). Washington, D.C.: U.S. Government Printing Office, Chapter 6, 227-309.
- Gilbreth, F. B. (1911). Motion Study. New York: D. Van Nostrand.
- Gilbreth, F. B. and Gilbreth, L. M. (1917). Applied motion study. New York: Sturgis and Walton.



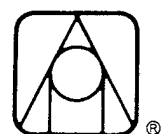
- Hale, H. B., McNee, R. C., Ellis, J. P., Bollinger, R. R., and Hartman, B. O. (1974). Endocrine-metabolic indices of aircrew workload: an analysis across studies. Proceedings of the AGARD Conference on Simulation and Study of High Workload Operations, AGARD-CP-146. (April) A10-1-A10-6.
- Harris, R., Glenn, F., Iavecchia, H., and Zaklad, A. (1986). Human operator simulator. In Karwowski, W., Ed., Trends in Ergonomics/Human Factors III, Part A. Proceedings of the Annual International Industrial Ergonomics and Safety Conference. Louisville, Ky. New York: North Holland.
- Hart, S. G. (1975). Time estimation as a secondary task to measure workload-attention sharing effect on operator performance. Proceedings of the Eleventh Conference on Manual Control. Moffett Field, CA: TMX-62-464, (May) 64-77.
- Hart, S. G., Battiste, V., and Lester, P. T. (1984). POPCORN: A supervisory control simulation for workload and performance research. Proceedings of the 20th Annual Conference on Manual Control. Washington, D.C.: National Aeronautics and Space Administration. NASA CR2341, pp. 431-454.
- Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall.
- Karwowski, W. and Ayoub, M. M. (1984). Fuzzy modeling of stresses in manual lifting tasks. Ergonomics, 27, 641-649.
- Knowles, W. B. (1963). Operator Loading Tasks. Human Factors, 5, 155-161.
- Kramer, A. F., Wickens, C. D., and Donchin, E. (1983). An analysis of the processing demands of a complex perpetual-motor task. Human Factors, 25, 597-622.
- Lane, N., Strieb, M. I., and Leyland, W. E. (1979). Modeling the human operator: applications to system cost effectiveness. In Modeling and Simulation of Avionics Systems and Command Control and Communication Systems. NATO/AGARD Proceedings No. CP-268.
- McCauley, M. E., Kennedy, R. S., and Bittner, A. C., Jr. (1980). Development of a performance evaluation test for environmental research (PETER): time estimation test. Perceptual and Motor Skills, 51, 655-665.
- McLucas, J. L., Drinkwater, F. J., and Leaf, H. W. (1981). Report of the President's Task Force on Aircraft Crew Complement. Prepared for the President, The White House. Washington, D.C.: U.S. Government Printing Office (July).
- Meister, D. (1985). Behavioral analysis and measurement methods. New York: John Wiley and Sons.



- Morey, N., Turkensen, B., Aidie, P., Drasic, D., Eisen, P., Kruschelnicky, E., Money, L., Schonert, H., and Thornton, C. (1986). Progress in mental workload measurement. Proceedings of the Human Factors Society 30th Annual Meeting. Santa Monica, CA: Human Factors Society, 1121-1123.
- Mostotsky, D. I. and Piedmont, R. L. (1985). Therapeutic Practice in Behavioral Medicine: A Selective Guide to Assessment, Treatment, Clinical Issues, and Therapies for Specific Disorders. San Monica, CA: Jossey-Bass.
- Mulder, G. (1973). Mental load and the measurement of heart-rate variability. Ergonomics, 16, 69-83.
- Norman, D., and Bobrow, D. (1975). On data-limited and resource-limited processes. Cognitive Psychology, 7, 44-64.
- Nunnally, J. C. (1967). Psychometric Theory. New York: McGraw-Hill.
- Reid, G. B., Shingledecker, C.A., and Eggemeier, F. T. (1981). Application of conjoint measurement to workload scale development. Proceedings of the Human Factors Society 25th Annual Meeting. Santa Monica, CA: Human Factors Society, 522-525.
- Rolfe, J. M. (1973). Whither workload. Applied Ergonomics, 4, 8-10.
- Roscoe, A. H. and Grieve, B. S. (1986). The impact of new technology on pilot workload. Warrendale, PA: Society of Automotive Engineers, Technical Paper No. 861773 (October).
- Ruggiero, F. T. and Fadden, D. M. (1982). Pilot subjective evaluation: a practical supplement to traditional assessment of workload. Proceedings of the International Conference on Cybernetics and Society. New York: Institute of Electrical and Electronics Engineers, 603-607.
- Sayers, B. M. (1973). Analysis of heart rate variability. Ergonomics, 16, 17-32.
- Seifert, O. J. and Chubb, G. P. (1978). SAINT: A combined simulation language for modeling large complex systems. Aerospace Medical Laboratory. Dayton, OH: Wright-Patterson Air Force Base, AMRL-TR-78-48.
- Senders, J. W. (1964). The human operator as a monitor and controller of multidegree of freedom systems. IEEE Transactions on Human Factors in Electronics, HFE-5, (September), 2-5.
- Senders, J. W., Kristofferson, A. B., Levinson, W. H., Dietrich, C. W., and Ward, J. L. (1967). The attentional demand of automobile driving. Highway Research Record, No. 195, 15-33.
- Shaffer, M. T., personal communication, November 19, 1986



- Shaffer, M. T., Shafer, J. B., and Kutche, G. B. (1986). Empirical workload and communications: analysis of scout helicopter exercises. Proceedings of the Human Factors Society, 30th Annual Meeting. Santa Monica, CA: The Human Factors Society, 628-632.
- Siegel, A. I. and Wolf, J. J. (1961). A technique for evaluating man-machine system designs. Human Factors, 3, p. 18.
- Siegel, A. I. and Wolf, J. J. (1969). Man-machine simulation models: psychosocial and performance interaction. New York: John Wiley and Son.
- Skipper, J. H., Rieger, C. A., and Wierwille, W. W. (1986). Evaluation of decision-tree rating scales for mental workload estimation. Ergonomics, 29 (4), 585-599.
- Stone, G., Gulick, R. K., and Gabriel, R. F. (1985). Use of task/timeline analysis to assess crew workload. Long Beach, CA: McDonnell Douglas Corp., Paper No. 7592.
- Sutcliffe, J. P. (1980). On the relationship of reliability to statistical power. Psychological Bulletin, 88, 509-515.
- Taylor, F. W. (1912). The present state of the art of industrial management. Transactions, ASME, 34.
- Tsang, P. S. and Wickens, C. D. (1984). The effects of task structures on time-sharing efficiency and resource allocation optimality. Proceedings of the Twentieth Annual Conference on Manual Control. Washington, D.C.: NASA-CP-2341, 305-317.
- Tulga, M. K. (1978). Dynamic decision making in multitask supervisory control: comparison of an optimal algorithm to human behavior. Cambridge, MA: MIT Man-Machine Systems Laboratory.
- U.S. Army Missile Research and Development Command (1979). Human Engineering Requirements for Military Systems, Equipment, and Facilities. MIL-H-46855B. Redstone Arsenal, AL: DRDMI-ESD.
- U.S. Army War College (1986). Army Command and Management: Theory and Practice. Carlisle Barracks, PA.
- Wierwille, W. W. (1979). Physiological measures of aircrew mental workload. Human Factors, 21 (5), 575-593.
- Wierwille, W. W. and Casali, J. G. (1983). A validated rating scale for global mental workload measurement application. Proceedings of the Human Factors Society 27th Annual Meeting. Santa Monica, CA: Human Factors Society, 129-133.



Wierwille, W. W., Casali, J. G., Connor, S. A., and Rahimi, M. (1985).
Evaluation of the sensitivity and intrusion of mental workload estimation
techniques. In W. B. Rouse, Ed., Advances in man-machine systems
research, Volume 2. Greenwich, CN: JAI Press, Inc., 51-127.

Wierwille, W. W., and Williges, R. C. (1978). Survey and analysis of operator
workload. Blacksburg VA: Systemetrics, Inc., S-78-101.

