Technical Report 997

Evaluating the Unit Performance Assessment System's After Action Review Displays

Theodore M. Shlechter and David W. Bessemer
U.S. Army Research Institute

Wade C. Rowatt and K. Paul Nesselroade, Jr.
University of Louisville

May 1994

United States Army Research Institute
for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Technical review by

Larry L. Meliza
Major Greg A. Walden

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POX, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1994, May	Final Jun 92 - Jul 93

4. TITLE AND SUBTITLE
Evaluating the Unit Performance Assessment System's After Action Review Displays

5. FUNDING NUMBERS
63007A
794
2124
H01

6. AUTHOR(S)
Shlechter, Theodore M.; Bessemer, David W.; Rowatt, Wade C.; & Nesselroade, Jr., K. Paul

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral and Social Sciences
Field Unit at Fort Knox
Fort Knox, KY 40121-5620

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
ATTN: PERI-IK
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
ARI Technical Report 997

11. SUPPLEMENTARY NOTES
Wade Rowatt and K. Paul Nesselroade, Jr., are graduate students at the University of Louisville working as consortium fellows at the U.S. Army Research Institute's Fort Knox Field Unit.

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

12b. DISTRIBUTION CODE
--

13. ABSTRACT *(Maximum 200 words)*

This research effort assessed military users' judgments about different versions of the Unit Performance Assessment System's (UPAS') performance feedback capabilities in after action reviews (AARs) for Simulation Networking (SIMNET) training. Interactive demonstrations with SIMNET training personnel and operational unit personnel were performed to determine the utility of the feedback materials generated by UPAS. Specific findings indicated that (a) the slide show method of exercise replay was preferred over an animated rerun method; (b) Battle Snapshot displays were more likely to be used than Exercise Timeline or Battle Flow displays; and (c) divergent opinions existed regarding the value of certain displays.

14. SUBJECT TERMS
Armor training
After action review
Performance assessment
Simulation networking

15. NUMBER OF PAGES
83

16. PRICE CODE
--

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Technical Report 997

Evaluating the Unit Performance Assessment System's After Action Review Displays

Theodore M. Schlechter and David W. Bessemer
U.S. Army Research Institute

Wade C. Rowatt and K. Paul Nesselroade, Jr.
University of Louisville

Field Unit at Fort Knox, Kentucky
Barbara A. Black, Chief

Training Systems Research Division
Jack H. Hiller, Director

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

May 1994

Army Project Number
2Q263007A794

Education and Training

FOREWORD

The Army has been continually challenged to provide effective training with efficient use of resources. The Simulation Networking (SIMNET) system was initially developed as a proof-of-principle technology demonstration of distributed interactive battle simulation. This technology has led to new concepts for affordable large-scale simulator training systems, such as the Close Combat Tactical Trainer (CCTT). It has also been designed to provide collective tactical training in a prototype training version, but with limited automated capabilities to support performance feedback in after action reviews (AARs). The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) has thus developed a low-cost automated system--the Unit Performance Assessment System (UPAS)--for producing AAR materials associated with SIMNET use.

This report describes research conducted to assess users' judgments regarding different versions of the UPAS materials. The feedback is being used to help refine these materials. Hence, the research and development process associated with UPAS involved a partnership among researchers, software developers, and training personnel.

The described research effort was part of the ARI Fort Knox Field Unit's work program entitled "Strategies for Training and Assessing Armor Commanders' Performance with Devices and Simulations (STRONGARM)." This task is supported by a Memorandum of Agreement entitled "The Effects of Simulators and Other Resources on Training Readiness," signed 16 January 1989. Parties to this agreement are the U.S. Army Training and Doctrine Command (TRADOC), the U.S. Army Armor Center and Fort Knox (USAARMC), the U.S. Army Materiel Command (AMC), and ARI.

The information contained in this technical report has been provided to training personnel and simulation training managers in the USAARMC. The research findings should be of general interest to training and simulation developers concerned with requirements for performance feedback.

EDGAR M. JOHNSON
Director

EVALUATING THE UNIT PERFORMANCE ASSESSMENT SYSTEM'S AFTER ACTION REVIEW DISPLAYS

EXECUTIVE SUMMARY

Requirement:

The Simulation Networking (SIMNET) technology for interactive battle simulation has been designed to provide collective tactical training for units from platoon to battalion level. The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) has developed the Unit Performance Assessment System (UPAS) to produce after action review (AAR) materials that may increase the training benefit of SIMNET exercises. This research was designed to assess users' judgments on the utility of the developed UPAS materials.

Procedure:

A series of preliminary and formal assessments was required to provide the information needed to determine UPAS' utility. The Preliminary Assessment Stage consisted of interviewing two Noncommissioned Officer (NCO) staff members from the Mounted Warfare Simulation Training Center (MWSTC) facility[1] and five Advanced Officer Basic Course (AOBC) instructors on the usefulness of and possible problems with the UPAS displays. The MWSTC and AOBC personnel examined computer displays produced by a partially developed UPAS package. The AOBC instructors also completed a brief questionnaire on the UPAS displays.

The formal assessment stage involved systematically assessing SIMNET instructors' ratings of the UPAS package. The participants for this stage were 30 instructors (15 officers and 15 NCOs) for the Command and Staff Department of the Armor School. Nine of the instructors were for the Armor Officer Advanced Course (AOAC) while six were for AOBC. The NCOs consisted of eight instructors for the Advanced Noncommissioned Officer Course (ANCOC) and seven for AOBC.

These instructors were shown two distinct types of Plan View Display methods--slide show and animated replay. They also saw the following UPAS outputs: Battle Snapshot, Exercise Timeline, Battle Flow, and graphic displays. They then completed a series

[1] This facility houses the SIMNET system.

of questionnaires on their preferences for using the different displays.

A validity check was then conducted to determine the usefulness of the Battle Flow, Battle Snapshot, Exercise Timeline, and graphic displays for actual AOBC training situations. Comments from five AOBC instructors were gathered as their students completed (a) platoon-level tactical road marches, (b) platoon-level offensive and defensive exercises, and (c) company-level offensive and defensive exercises.

A generalizability check regarding UPAS' potential for unit training purposes was also conducted. Seven company-grade officers from operational units completed a brief questionnaire on the displays' usefulness.

Findings:

The findings indicated that SIMNET training personnel would like to have the different UPAS displays available, including both methods of presenting the Plan View Display replay, for use in their AARs. Also, the slide show method was found to be preferable to the animated method, and the Battle Snapshot displays were likely to be used more than the Exercise Timeline and Battle Flow.

Differences were found between officers and NCOs regarding the value of some of these displays. The AOBC NCOs tended to have the most favorable view of the two Plan View Display methods while the AOBC officers tended to have the least favorable view. Also, the AOBC and ANCOC NCOs tended to be more favorable toward the Exercise Timeline and Battle Flow than were the AOBC and AOAC officers.

Utilization of Findings:

These findings will be used by the U.S. Army Armor Center and the Simulation Training and Instrumentation Command to support decisions on further improvement and fielding of the UPAS. They will also be used to help determine the feedback requirements and capabilities for the Close Combat Tactical Trainer.

EVALUATING THE UNIT PERFORMANCE ASSESSMENT SYSTEM'S AFTER ACTION REVIEW DISPLAYS

CONTENTS

CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

CONTENTS (Continued)

CONTENTS (Continued)

EVALUATING THE UNIT PERFORMANCE ASSESSMENT SYSTEM'S AFTER ACTION REVIEW DISPLAYS

Introduction

This report focuses on evaluation of the Unit Performance Assessment System's (UPAS') package of feedback--after action review (AAR)--displays to be used with the Simulation Networking (SIMNET)[1] system. This report also describes the partnership among researchers, software developers, and military training personnel in refining UPAS.

Need for SIMNET

Contemporary military trainers have been challenged to provide more effective training with dwindling financial resources. The Army has found that training officers in tactical skills in a field environment has become increasingly expensive. However, inexpensive table exercises and board games have not faithfully reproduced the conditions inherent in field exercises (Kristiansen, 1987). Needs for economical means of collective training have been identified in Defense Science Board reports (e.g., Defense Science Board, 1976, 1988).

The Defense Advanced Research Projects Agency developed the SIMNET system as a proof-of-principle technology demonstration of real-time distributed interactive battle simulation (Alluisi, 1991). It has also been designed to provide collective tactical training for units from platoon to battalion levels. SIMNET consists of the integrated use of training simulators, combat support equipment, and instructor's stations. Each simulator has its own microcomputer and is connected with the other simulators by specifically formatted data packets across the computer network (Garvey & Radgowski, 1988). These data packets contain information about vehicle appearance (e.g., location, speed, and type of vehicle), vehicle status (e.g., fuel and ammo level of tank), direct fires (e.g., hits, type of target, and rounds fired), indirect fires (e.g., location and results of artillery fire), impact (results of vehicle firing), and status change (destruction or damage) of the vehicle. Each simulator's computer independently uses the information to construct views of the simulated battlefield and to react to events.

The MWSTC is used for the following Armor School courses: (a) the Armor Officer Basic Course (AOBC), (b) Armor Officer Advanced Course (AOAC), and (c) Armor Noncommissioned Officer Course (ANCOC). These courses use this facility as part of their training in tactics. Various operational units often use the

---

[1] In this report, the term SIMNET refers to the training system. The term Mounted Warfare Simulation Training Center (MWSTC) refers to the facility that houses the SIMNET system.

MWSTC as part of their sustainment training. Armor and mechanized infantry units may also use the MWSTC as part of their preparation for the National Training Center.

Despite some methodological flaws, the results of several evaluations have demonstrated SIMNET's effectiveness for some armor training purposes (Bessemer, 1991; Brown, Pishel, & Southard, 1988; TEXCOM, 1990; Shlechter, Bessemer, & Kolosh, 1991). Shlechter et al., for example, found that SIMNET training provided armor students with the needed practice opportunities to develop their tactical skills. Bessemer (1991) suggested that SIMNET's effectiveness was also related to the feedback provided by instructors. He noted that SIMNET's effectiveness appeared to increase during the course of his evaluation as improvements were made to the corresponding AARs. The instructors appeared to more effectively use the SIMNET system as they gained experience with using it. Improvements in these AARs also may be attributed in part to additional SIMNET equipment--e.g., the stealth vehicle. The stealth vehicle provides a direct view of the battleground from an invisible vehicle moving on or above the terrain. This added capability made it easier for training personnel to observe the SIMNET exercises and subsequently provide feedback to their students.

Need for Developing UPAS

Although an exercise replay capability has been provided, SIMNET's feedback capabilities are still limited. Although all exercises on the network are recorded, only one platoon or company can be observed effectively at one time during playback. Furthermore, the replay system is not portable, which means that the training personnel must use it in a specified area. Therefore, the system does not support replays for several units conducting AARs simultaneously at different locations in the SIMNET. Those training personnel who do not have access to this system must then sketch their unit's course of action on a situational map; the fidelity of such sketches, however, depends upon the ability of the training personnel and the unit personnel to note or accurately remember important incidents.

Another limitation is that the SIMNET sites do not have the capability to provide graphic or tabular summaries of quantitative measures reflecting the unit's performance. Such materials would supplement the existing replay system by providing more precise information about particular aspects of a unit's performance. Such summaries would help training personnel to clearly elucidate bases for many of the salient lessons learned from the exercises.

Easter, et al. (1986) suggested that an automated performance measurement system with graphic replay capabilities could help eliminate such limitations with a training system's (e.g., SIMNET) feedback capabilities. An automated performance measurement system would provide students and instructors with

timely and useful feedback by performing all statistical analyses in real or near-real time (Easter, et al.). An automated performance measurement system would also provide printed reports of this feedback and monitor every significant aspect of a unit's SIMNET performance.

UPAS is a microcomputer system that collects and records the previously discussed data packets from SIMNET. UPAS also translates and organizes the derived information into a relational database (Meliza, Bessemer, Burnside, & Shlechter, 1992). From the packet file and database, the collected information can be further manipulated into map and graphic displays of unit performance that can be used during the SIMNET AARs. UPAS can also print these displays for later use.

Questions remain, however, about the most appropriate format(s) for these AAR displays. The educational and training literature does provide some limited guidance in developing these formats (see Downs, Johnson, & Fallesen, 1987; Garlinger, 1987; Holding, 1965; Kulhavy & Stock, 1989; Meliza, Bessemer, et al., & Shlechter, 1992; Pridemore, Webb, Haygood, Stock, & Kulhavy, 1990; Scott, 1983). Kulhavy and Stock suggested that such feedback materials must contain enough information to correct students' faulty perceptions of their performance without overwhelming them. Meliza, Bessemer, et al. suggested that feedback materials for collective armor training exercises must include concrete examples of problematic unit performance for crucial mission events.

The Prototype UPAS Displays

Meliza, Bessemer, et al. (1992) developed a series of prototype AAR displays for SIMNET linked to the UPAS database. These prototype displays are: Battle Flow, Battle Scorecard, Battle Snapshot, Exercise Timeline, and Plan View Display. The R & D team has developed the Battle Flow and Scorecard based on guidance by SIMNET managers having long experience in unit training; the Battle Snapshot and Plan View Display were based on displays used at the National Training Center (NTC), also at the suggestion of SIMNET managers; the Exercise Timeline was developed based on ideas from a member of the R & D team.

The following descriptions of these materials have been based upon information presented in the SIMNET UPAS User's Guide and accompanying technical reports (Meliza, Tan, White, Gross, & McMeel, 1992; Meliza, Bessemer, et al., 1992). The Exercise Timeline and Battle Scorecard provide, respectively, graphic and tabular overviews of the unit's performance. As shown in Figure 1, the Battle Scorecard presents quantitative information about the number of hits, kills, and misses for both sides. Such information can provide the unit with a picture of the mission's success.

DIRECT FIRE WEAPON SYSTEM SUMMARY

SIDE	FIRING WEAPON	FIRING AMMO	SHOTS	MBT_H	MBT_K	IFV_H	IFV_K	OTH_H	OTH_K	TOT_H	TOT_K
B	US M1	US M392A2 - 105mm/KP	95	4	4	0	1	0	0	4	5
		US M456A1 - 105mm/SCP	87	0	0	6	5	0	0	6	5
Total			182	4	4	6	6	0	0	10	10
SIDE	FIRING WEAPON	FIRING AMMO	SHOTS	MBT_H	MBT_K	IFV_H	IFV_K	OTH_H	OTH_K	TOT_H	TOT_K
B	US M2	US M791 - 25mm	197	2	0	1	3	0	0	3	3
Total			197	2	0	1	3	0	0	3	3
SIDE	FIRING WEAPON	FIRING AMMO	SHOTS	MBT_H	MBT_K	IFV_H	IFV_K	OTH_H	OTH_K	TOT_H	TOT_K
R	USSR T72M	US M392A2 - 105mm/KP	1	0	1	0	0	0	0	0	1
		US M456A1 - 105mm/SCP	1	0	0	0	0	0	0	0	0
Total			2	0	1	0	0	0	0	0	1
SIDE	FIRING WEAPON	FIRING AMMO	SHOTS	MBT_H	MBT_K	IFV_H	IFV_K	OTH_H	OTH_K	TOT_H	TOT_K
R	USSR_BMP2	US M791 - 25mm	155	3	0	0	0	0	0	3	0
Total			155	3	0	0	0	0	0	3	0
Overall total			536	9	5	7	9	0	0	16	14

Figure 1. Sample of the Battle Scorecard Display.

The Exercise Timeline provides information regarding the temporal occurrences for crucial mission events, e.g., crossing operational control lines, and first friendly and enemy fire (Figure 2). The timeline shows occurrences of types of communications as recorded by an observer. The Exercise Timeline can also help the instructor to use the other feedback materials more effectively. For example, the time scale can help the SIMNET instructor choose the most appropriate periods of activity to be replayed by the Plan View Display.

As shown in Figures 3-5, the Plan View Display, Battle Flow and Battle Snapshot graphically show activities on the SIMNET terrain map. These displays represent major terrain features and grid coordinates along with symbolic icons for different vehicles. The Plan View Display plays back either the entire exercise or segments of the exercise. The Battle Flow provides a line trace of the unit's movements across time increments. The Battle Snapshot shows the unit's position for salient moments of the exercise.

Based on content analysis of tasks trainable by SIMNET, Meliza, et al. (1992) has concluded that each display would make a unique contribution to the SIMNET AARs. The Battle Snapshot, for example, would provide the best picture of vehicle orientation during the battle. Possibly then, the UPAS displays can make an important contribution to different training programs which use MWSTC.

Need for Instructional Personnel in the R & D Process

Questions remain about the potential utilization of the UPAS displays. The intended instructional personnel may feel that these displays are not suitable for their training purposes. Or, they may not use the UPAS materials because of some factors which are not readily apparent to the R & D team.

A major theme in the history of instructional technology has been an initial widespread enthusiasm for an instructional innovation followed by either its limited use or eventual disuse (Cuban, 1986; Shlechter, 1991). Film, radio, television, and computer-based instruction were hailed by their generation of instructional developers and policy makers as the educational panaceas of their day (Cuban; Shlechter). And yet, these innovations have never been widely accepted by instructors as instructional delivery systems.

Cuban (1986) and Thomas (1987) noted two main reasons for this historical trend. One, instructors have resisted changes which are seen as imposed solutions. Cuban, for example, argued that educational television was developed and implemented in school districts by non-teachers without soliciting the advice or consent of the user--teachers. Two, instructors have been hesitant to employ any instructional technologies (e.g., computer-based delivery systems) which are difficult for them to

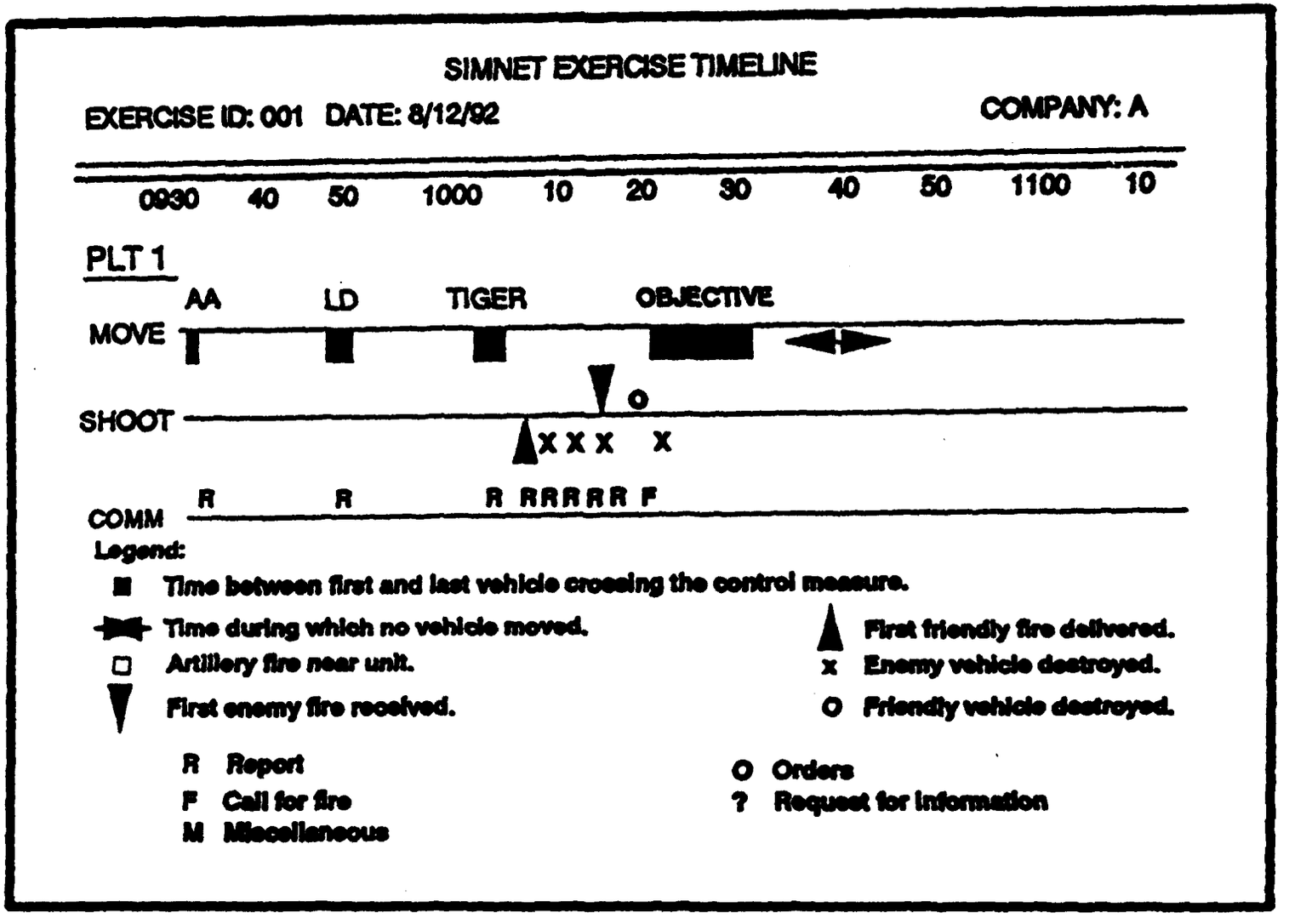


Figure 2. Sample of the Exercise Timeline Display.

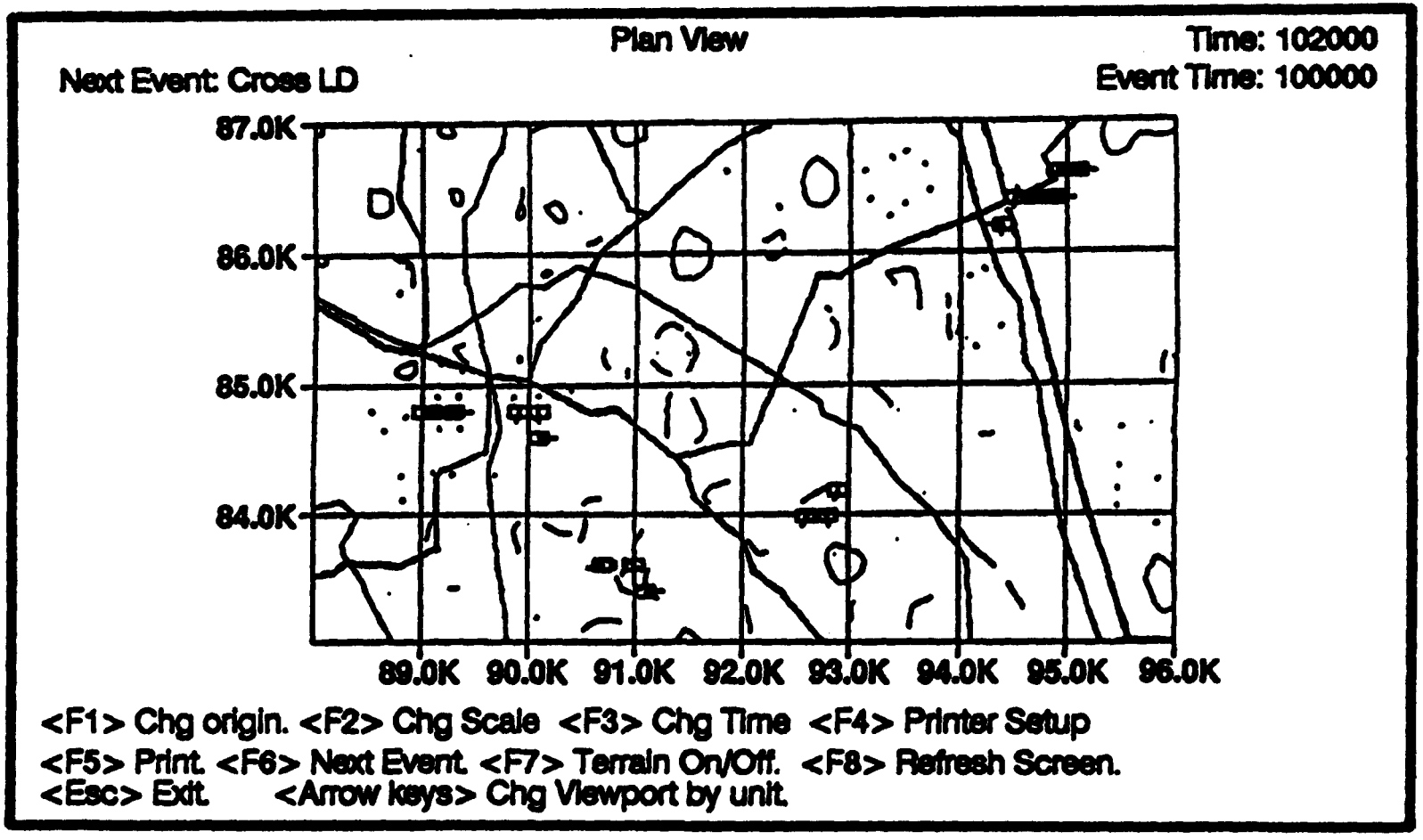


Figure 3. Sample of the Plan View Display.

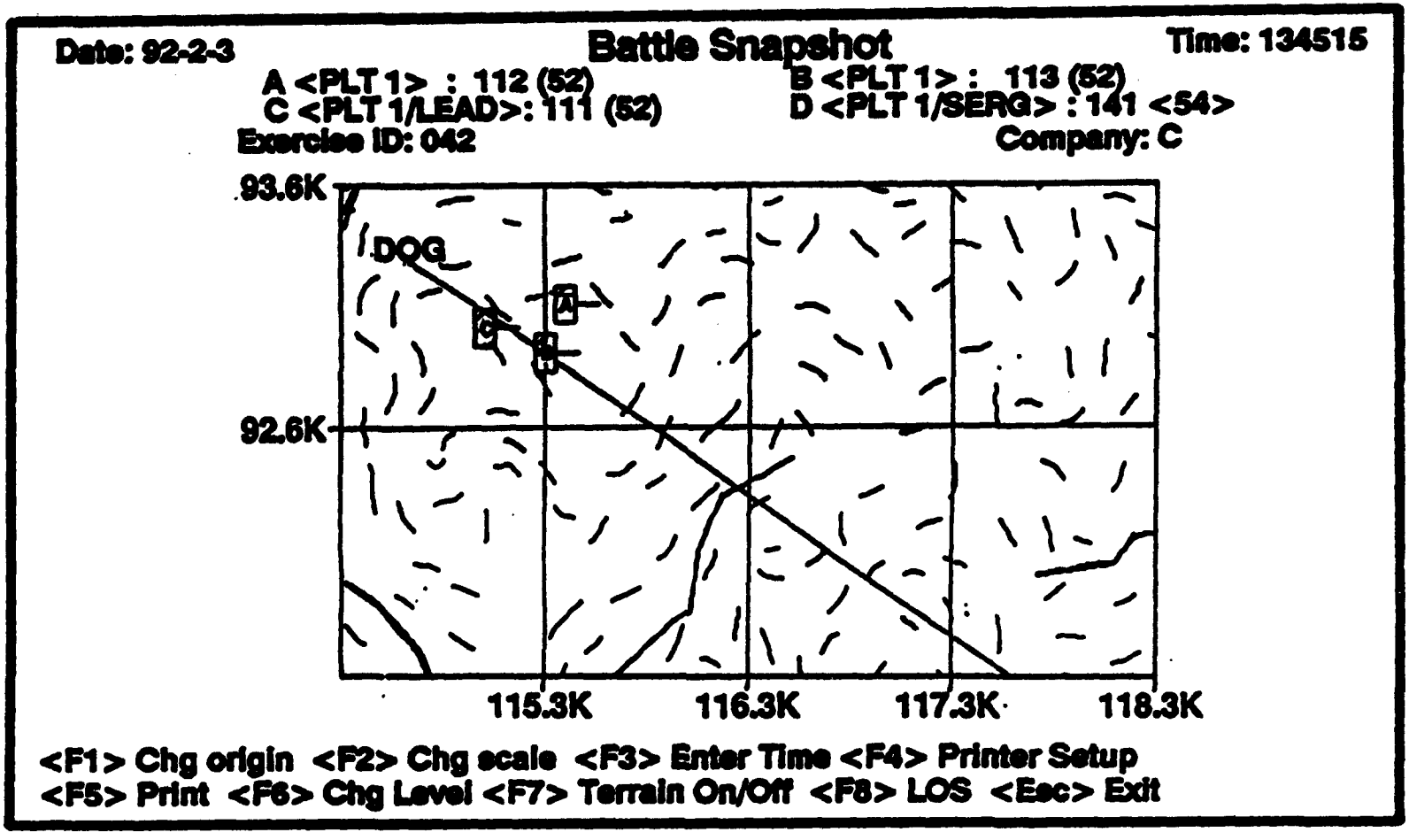


Figure 4. Sample of the Battle Snapshot Display.

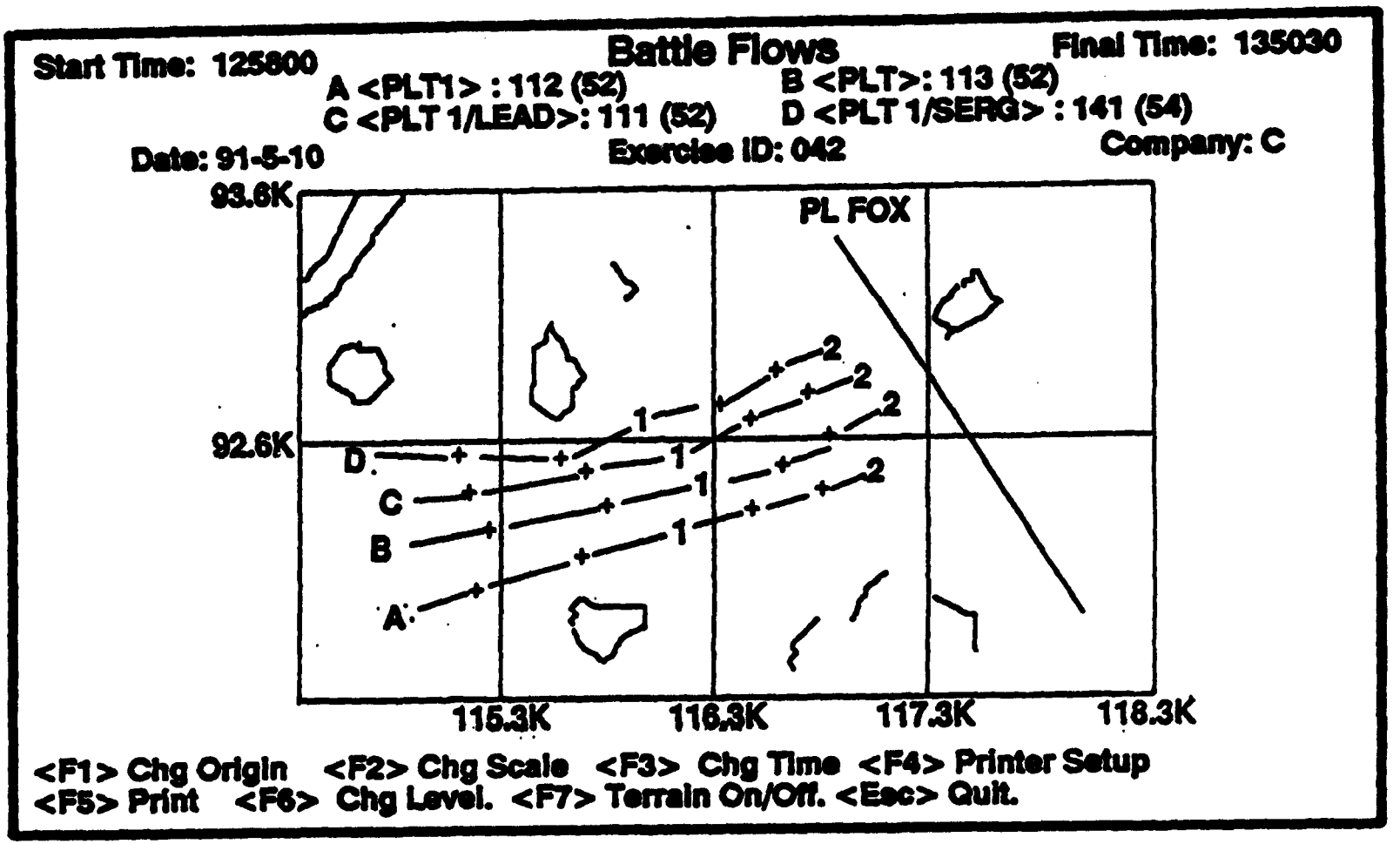


Figure 5. Sample of the Battle Flowchart Display.

use (Thomas). Hence, instructors' sense of ownership and ease with using the product are key factors in its utilization.

The military has also found that these two factors underlie instructors' acceptance of a new system (Gray, Roberts-Gray, & Gray, 1983; Polzella, Hubbard, Brown, & McLean, 1987; Reidel, 1988; Roberts-Gray, 1983; Shlechter, Burnside, & Thomas, 1987). Shlechter, et al. reported that military instructional personnel must assume "ownership" of the newly developed instructional program; otherwise, the program will never be fully utilized.

Gray, et al. (1983) maintained that communications problems between training personnel and instructional developers might make a high quality product very difficult to use, a situation which widens the gap between development and utilization of instructional innovations. They have developed a framework for implementing an instructional innovation that involves: (a) analyzing the potential implementation problems, (b) planning the implementation process, (c) executing the implementation program, and (d) monitoring and providing feedback to the user regarding the implementation process. At each stage, the developer should work with the potential users in order to increase the latter's ability to accept, operate and manage the new training product (Gray et al.). This framework, however, does not deal with the R & D involved in producing the instructional innovation.

Involving instructional personnel with the R & D phases would ensure their sense of "ownership" of the instructional innovation. It would correspondingly reduce the possibility of wasting precious funds on developing instructional products that are either never fully utilized or need costly fixes. Instructors may also not be very accepting of a product which takes too much time to eicher operate or provide the needed information. Military instructors, for example, may not be willing to wait very long for an automated feedback system to provide the materials needed for their SIMNET AARs. R & D teams, however, rarely consider the effects of such temporal factors on instructors' willingness to embrace a new innovation. Perhaps then, R & D teams working with the intended instructional personnel could produce instructional innovations which do not take too much time away from the instructional program; otherwise, the innovation will probably not be utilized.

Research Objective

This research effort was designed to obtain feedback from potential UPAS users regarding the training value of the developed AAR displays. The feedback obtained was used to help make any enhancements to the AAR displays.[2]

---

[2] Difficulties encountered during the preparation of display examples (the stimulus materials) were also noted and used to recommend user interface modifications.

SIMNET instructors have previously indicated that the Battle Scorecard, Battle Flow, Battle Snapshots, and Exercise Timeline prototypes are instructionally suitable for their training purposes (Shlechter, Meliza, Bessemer, & Burnside, 1992). These findings, however, were based on paper drawings of the display concepts, which may not fully represent the developed displays.

Preliminary Assessment Stage

Method

Two MWSTC staff members and five AOBC instructors were interviewed concerning their judgements regarding a partially developed set of AAR displays--the Exercise Timeline was not fully operational. The MWSTC staff members were non-commissioned officers (NCOs) while the AOBC instructors were officers. The MWSTC staff members and AOBC instructors were interviewed while viewing the displays with their colleagues. They were shown a series of display examples, which were selected to demonstrate the variety of capabilities provided by UPAS.

The interviews, which required over 90 minutes per group, addressed the following training concerns:

1. usefulness of the different displays for either a specific training program or for any training program which uses SIMNET (Instructors were to assume that each display was easy to use and fully operational.)

2. problems with understanding or using the displays.

3. enhancements needed to make the particular display more instructionally suitable.

4. general enhancements needed to UPAS.

5. reasons for their responses to issues 1-4.

The AOBC instructors also completed a brief questionnaire. As shown in Appendix A, this questionnaire consisted of having the instructors rate the usability of each display from not usable to indispensable on a scale developed by Polzella and Hubbard (1986). This questionnaire also asked the instructors to make any additional comments about the system. The MWSTC staff members did not have the time to complete this questionnaire.

Findings

Three similar concerns were noted during the group interviews with the MWSTC staff members and AOBC instructors. One, the MWSTC staff members and AOBC instructors exhibited concern about the time delays associated with accessing the data. They wanted immediate access to the data. As noted by one AOBC instructor:

> Currently, it (the Plan View Display) would take too long to analyze and glean pertinent data to be of much use.

Two, they indicated a need for a better vehicle identification scheme. The MWSTC staff members suggested that these icons be labeled by the standard numbering system used for such purposes--e.g., 1 for platoon leader, 4 for platoon sergeant, and 2 and 3 for wingmen--rather than by the letter scheme used in the Battle Snapshot and Battle Flow displays. Four of the five AOBC instructors elaborated upon this issue in their questionnaire responses with two indicating a requirement for bumper numbers and two for the scheme suggested by the MWSTC staff members.

The AOBC instructors also suggested during the interviews that this package include fire fight and communication displays. They suggested that the fire fight display should include color coded vectors indicating hits, kills, and misses for each tank in the unit. This display should also include icons indicating positions of: (a) live vehicles for either their final positions or a set time, and (b) dead vehicles. These instructors also claimed that the Battle Snapshot needed a fan line-of-sight vector with the selected tank's line-of-sight and "blind-spots" being color-coded. Two of the AOBC instructors elaborated upon these points in their questionnaire responses.

Three, a need was expressed to make the display information more visible to the students. The MWSTC staff members wanted to know if it was possible to project the PVD information on the SIMNET stealth screen. Three of the AOBC instructors noted on their questionnaires that UPAS needs the ability to print screen information.

Divergent viewpoints were expressed about the usability of the Battle Scorecard and Exercise Timeline. The AOBC instructors indicated on the questionnaires that the Exercise Timeline was fairly to extremely useful, but felt that the Battle Scorecard had limited value (see Figure 6). They felt that graphs depicting relationships of rounds fired over time and rounds fired over range were more appropriate for their purposes than was the Battle Scorecard. Both of the MWSTC staff members liked the Battle Scorecard as it reflected information presented at the National Training Center.

The MWSTC staff members, however, had problems with the Exercise Timeline. They thought that this display was of limited value. It was also difficult for them to distinguish the activities attributable to the friendly or oppositional forces. These concerns might not have been voiced if a fully operational Exercise Timeline had been available.

Based on the aforementioned comments, changes were made to the displays. A function to provide line-of-sight vectors was

added to the Battle Snapshot display. Communication times were added to the TImeline display. A slide show Plan View Display method was also developed as a response to the instructors' concerns about the current Plan View Display. A slide show method would eliminate the discussed time delays by using a software program that allows the UPAS system to store and automatically, "freeze-frame" screens of the SIMNET action in graphic-formatted files. The idea to freeze-frame the PVD screen came from an AOBC instructor. The files could then be shown in a fixed order by a presentation program.

Formal Assessment Stage

The formal assessment had three objectives. One, to systematically examine SIMNET users' views on the prototype UPAS displays. A major focus of this examination was the users' views on the animated rerun and slide show methods of presenting Plan View Display exercise replays. Also, the judgements of a broader sample of users on the desirability and usefulness of the other UPAS displays was needed to confirm the previous findings. Two, to explore the underlying reasons for their preference, if any, for the animated rerun or slide show methods. Three, to determine if systematic differences exist among different types of SIMNET users regarding the value of the UPAS' displays.

Method

Participants. Thirty instructors (15 officers and 15 NCOs) from the Command and Staff Department of the Armor School were the participants. Nine of these officers were instructors for the Armor Officer Advanced Course (AOAC) while six were instructors for AOBC. The NCOs consisted of eight instructors for the Advanced Non-Commissioned Officer Course (ANCOC)[3] and seven for AOBC.

These Armor School instructors had spent the following amounts of time training students at the MWSTC: a) AOAC officers--11.4 mos.; b) AOBC officers--12.2 mos.; c) ANCOC NCOs--23.1 mos.; and d) AOBC NCOs--16.7 mos. Also, most of them had prior SIMNET and NTC experience. One-way analysis of variance tests failed to reveal any significant differences among the four types of instructors for these demographic variables. Other demographic data are presented below in Table 1.

Materials. A set of Plan View Display replays was prepared to illustrate the following types of mission exercises in MWSTC: (a) platoon-level tactical road march (TRM), (b) platoon-level force-on-force (FOF) exercise, (c) company-level attack (CL ATK) exercise, and company-level defense (CL DEF) exercise. All exercises, except the FOF exercises, employed semi-automated force vehicles as the opponent force. The replays used UPAS data

[3] The ANCOC course is only taught by NCOs.

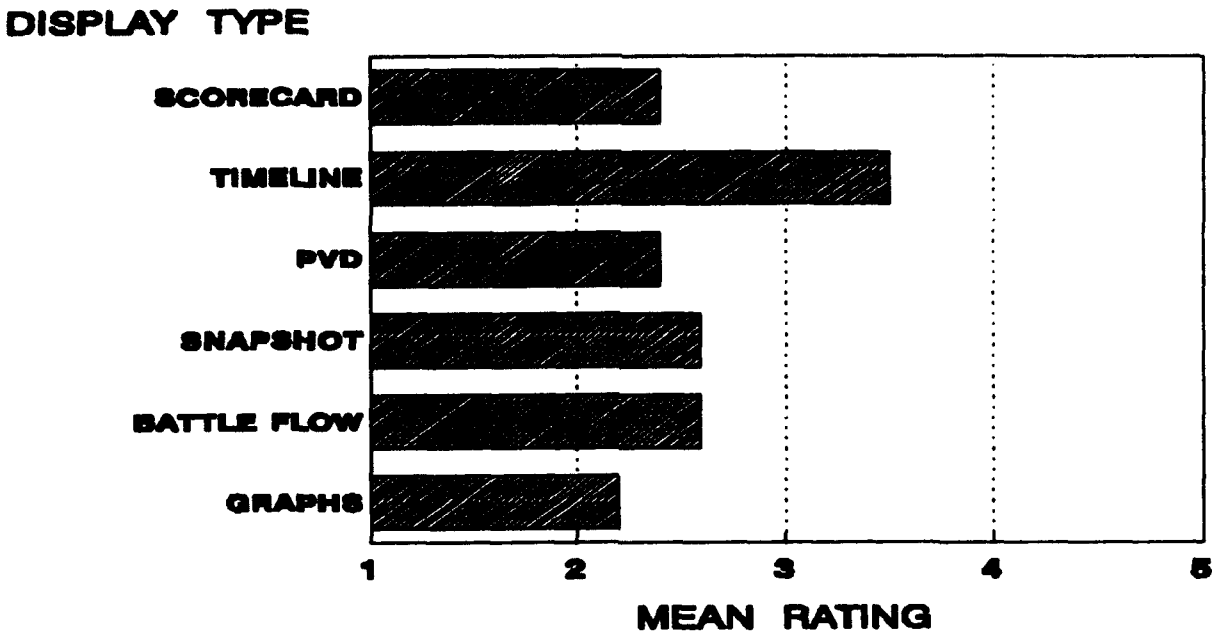


Figure 6. Usefulness of UPAS AAR aids rated by Armor School instructors. Responses on the 5-point rating scale were: "not useful" (1), "slightly useful" (2), "fairly useful" (3), "extremely useful" (4), and "indispensable" (5).

files recorded from actual exercises conducted in an AOBC class.

Table 1

Selected Demographic Variables for the Armor School Instructors

	OFFICERS		NCOs	
Background Characteristic	AOAC (n=9)	AOBC (n=6)	ANCOC (n=8)	AOBC (n=7)
Prior SIMNET Experience*	8	3	3	3
Prior NTC Experience*	9	5	4	3
Desert Storm Experience	2	2	2	5

*Significant difference ($p < .05$) between Officer and NCO groups.

The set of eight Plan View Display replays included two versions for each type of mission. The first version demonstrated the animated (free-running) rerun method, and the second version demonstrated the slide show method. For the animated rerun method, an operator's script indicated the display

settings to be used and temporal intervals to be replayed. A sample of coinciding views was captured for presentation by the slide show method. The views were taken from the Plan View Display at irregular intervals averaging 1-2 minutes apart, at time points that showed key events. Four presentations of selected examples of the Battle Flow, Battle Snapshot with/without line-of-sight, Exercise Timeline and graph displays were also developed as slide shows for the same exercises. Software faults prevented the use of Battle Scorecard displays.

The evaluation materials also consisted of three attitudinal questionnaires (A-C as shown in Appendices B-D, respectively). These questionnaires consisted of: (a) Likert-scaled preference items, (b) forced choice preference items, (c) continuous rating scale items, and (d) open-ended response items. The 5-point response scale for the Likert items was taken from the Questionnaire Construction Manual (Babbitt & Nystrom, 1989). Other forced choice and rating items used scaled adjectives selected from the same source. The rationale behind developing these questionnaires is addressed in the section on criterion measures.

Procedure. Each instructor viewed, in groups of five or fewer, the Plan View Display replays and examples of other UPAS displays during one 2-hour session. In the first hour, the instructors: (a) saw each replay method being demonstrated for two of four missions, (b) completed Questionnaire A after each demonstration, and (c) completed Questionnaire B after viewing all four replays. Half of the instructors saw the methods in the following sequence: animated rerun, slide show, slide show, animated rerun. The remaining instructors saw the methods in an opposite sequence: slide show, animated rerun, animated rerun, slide show. During the second hour, the instructors viewed the set of Battle Flow, Battle Snapshot, Exercise Timeline and graph materials and then completed Questionnaire C.

Several days after this viewing session, the AOAC instructors were interviewed about possible UPAS improvements. This group interview lasted approximately 30 minutes. Because of their duty requirements, the AOB and ANCOC instructors were not available to be interviewed at this later date. And, the prolonged evaluation session prohibited conducting such interviews immediately after the instructors completed Questionnaire C.

Criterion Measures. As indicated, the criterion measures consisted of the instructors' responses to the questionnaires. Each questionnaire was designed to assess a different facet of the previously discussed research objectives. The instructors' preferences for the two replay methods were assessed by their responses to Parts A and C of Questionnaire B. Having seen two examples of each method, the instructors possessed the information needed to rate the animated rerun and slide show replays.

Items 1-5 of Part A and one item in Part C examined the instructors' preference for a particular replay method. Item 1 called for the direct choice of one, both, or neither method for use in AARs. Items 2 and 3 asked the instructors to compare each replay method to what they do now in AARs. The 7-point scale for these items ranged from "much worse" to "much better." Item 4 and the Part C item both asked which method they preferred, using a 7-point scale between "much prefer (the animated rerun)" to "much prefer (the slide show)." These identical items enabled any change in preference to be detected, comparing answers before and again after other items drew attention to particular aspects of each replay method, and to the time factors involved in AARs. Item 5 asked the instructors to estimate on a continuous scale of 0-100% of AAR's, how often they would use each replay method.

Reasons for these preferences were assessed by analyzing the instructors' responses to Questionnaire A and Items 6-13 of Part B. Questionnaire A probed the instructors' feelings on the animated rerun's and slide show's instructional characteristics in relationship to the AAR process. The statements were worded to be equally applicable to both methods. Also, repeated responses to the same items for replays presented by different methods provided an indirect means of comparing ratings between methods. The first three items contained statements about the replay's help to the exercise instructors in an AAR. These items were based on the focal points of an AAR--what happened, why it happened, and how to improve in the future. The fourth item concerned the replay's help to the trainer in identifying training needs. The fifth item called for an overall evaluation, asserting that the replay would "help improve" AAR effectiveness. Responses to this item would also provide insights into the instructors' feelings on using the slide show and animated rerun replays. The last item addressed usage of time in AARs, stating that the replay would "take too much time."

Items 6-13 on Questionnaire B assessed the instructors' feelings toward specific characteristics of the animated rerun and slide show methods. These characteristics were operating: (a) procedures (Questions 6 and 7), (b) time (Questions 8 and 9), and (c) control (Questions 10 and 11). Also assessed were the instructors' feelings on the flow of information associated with the two methods (Questions 12 and 13).

Questionnaire C, Part A examined instructors' judgements about the Battle Flow, Battle Snapshot with/without line-of-sight, Exercise Timeline and graphic displays. Instructors completed the questionnaire after presentations of examples of each type of display for the same four missions shown in the replays. Item 1 probed the instructors' preferences for which display(s) should be kept in the UPAS' package (inclusion responses). This item also assessed the instructors' rankings on the order of priority (rank 1 = highest to rank 5 = lowest) for retaining each display. Items 2-6 assessed the usefulness of the different displays for different types of AOBC missions. Each

display was judged for four missions using a 5-point scale that ranged from "not useful" to "very useful."

Part B in both Questionnaires B and C contained scaled items on the time available before and during AARs for presenting the UPAS materials. A detailed discussion of these results will be presented in a forthcoming report on the temporal issues associated with operating UPAS.

The questionnaires also contained room for comments about UPAS from these instructors. They were asked to:

1. indicate reasons for their responses to each item on Questionnaire A and items 1-6 on Questionnaire B;

2. make any suggestions about ways of improving the slide show or animated rerun on Questionnaire B;

3. make any final comments about UPAS.

Data Analyses. Criteria measures associated with the responses on ordered category scales or continuous scales were numerically coded. This coding was done so that linear model analyses of variance could be performed on the resulting scores. Multivariate repeated measures analyses compared the average responses to groups of conceptually related items (e.g, items on instructional characteristics), and univariate analyses were done on single unrelated items. Comparisons addressing main issues (e.g., comparisons between Plan View Display methods) were tested for significance with $\alpha = .05$. Tests involving variables of secondary interest (e..g, interaction between Groups and Displays) were performed using $\alpha = (.01)df$, where df is the degrees of freedom associated with the tested hypothesis. This choice of α levels was based on a suggestion of Kepple (1982). (See Appendices E-G for a more technical description of these data analyses and results.)

Content analyses were performed on the instructors' comments regarding suggested improvements to the UPAS displays. Two judges working together determine the appropriate categories for these comments. One of these judges then tabulated comments per category.

Results and Discussion

The results are discussed in relationship to the instructors': (a) preferences for the animated rerun and slide show replays (Plan View Display preference results), (b) ratings of the instructional characteristics of the two replay methods (instructional characteristics results), and (c) judgements about the Battle Flow, Battle Snapshot, Exercise Timeline, and graph displays (supporting display results). Instructors comments

about improving the UPAS package are also presented and discussed.

Plan View Display Preference Results. A significant number of instructors (63%) indicated a preference for having both methods available for their AARs (see Figure 7). This figure also shows that 13% wanted neither method, 13% wanted slide show only, and 10% wanted animated rerun only. Instructors who wanted to use both methods also claimed that these methods would be utilized for approximately 60% of their AARs (see Figure 8). Those instructors who preferred a particular method would utilize this method for approximately 80% of their AARs.

Also, a significant number of the instructors favored the slide show over the animated rerun when forced to choose between the two displays--57% for the slide show method and 24% for the animated rerun method (see Figure 9). This preference pattern was manifested for the question that requested the instructors to carefully consider the different instructional features and time factors associated with using these methods (Question C of Part B). Significant differences were not found for Question 4 of Part A, which did not ask the instructors to consider these issues. For this question, 43% chose the slide show method, 34% chose the animated rerun method, and 20% indicated both methods. The change in response pattern between these questions was significant.

The preference data have thus shown that instructors would utilize both replay methods with the slide show method being the preferred method. Their preference for the slide show method was seemingly a function of its relative instructional effectiveness. (The reader is referred to Appendix F for a further description of these results)

Instructional Characteristics Results. A significant difference favoring the slide show replay was found for the statement--"Using this display would help improve the effectiveness of an AAR." As shown in Figure 10, 61% concurred with this statement for the slide show method but only 45% for the animated rerun method. Also, 56% of the instructors agreed that the slide show method would be better for helping students pinpoint their problems while 41% agreed with this statement for the animated rerun method (see Figure 11). These differences were statistically significant. The data analyses also revealed a trend favoring the slide show method for the statement--"This replay in the AAR would help the unit find ways to improve their mission performance." Fifty eight percent of the instructors agreed with this statement for the slide show method as compared to 51% for the animated rerun method (see Figure 12). The previously found preference for the slide show method was thus based on its relative instructional effectiveness.

Temporal considerations also influenced preferences for the slide show method. A significant difference favoring the slide

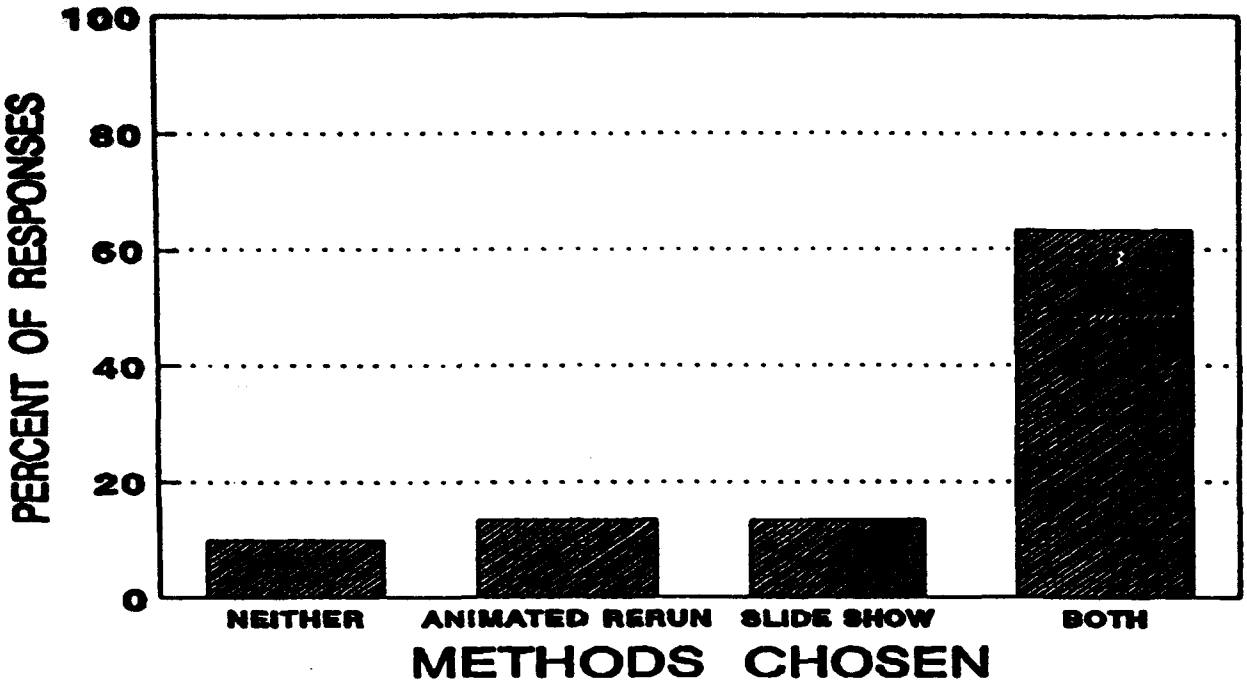


Figure 7. Relative choice between SIMNET exercise replay methods that Armor School instructors wanted in UPAS for use in AARs.

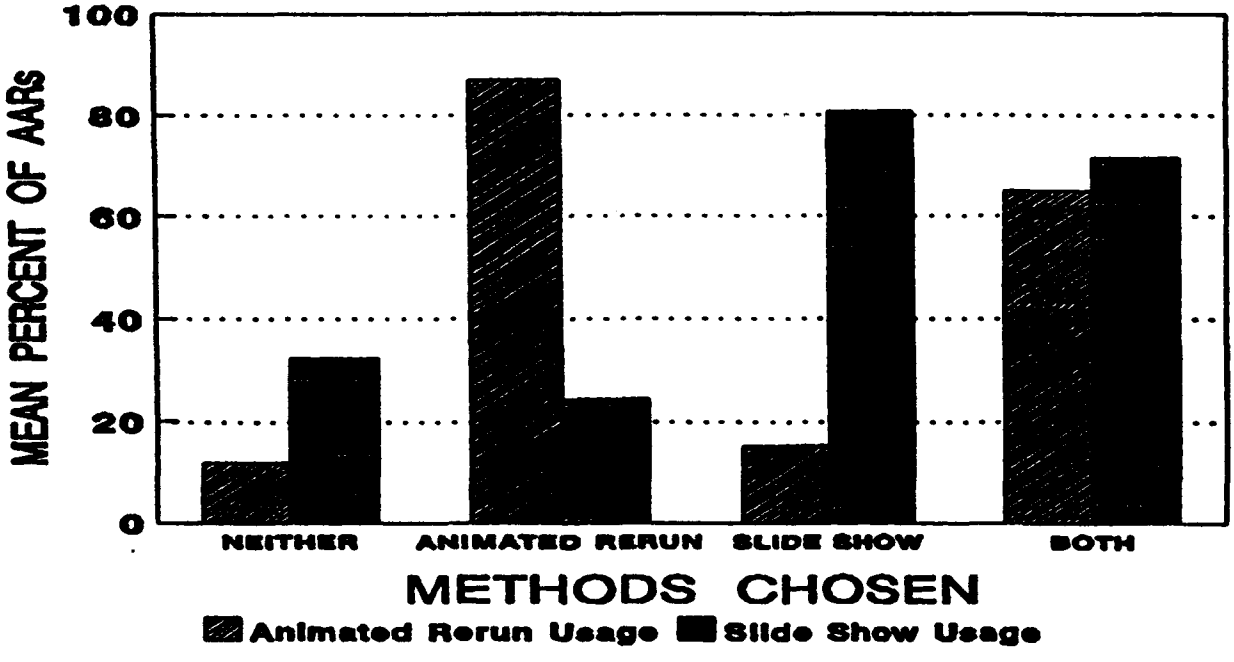


Figure 8. Estimated usage of exercise replay methods in AARs by instructor groups that chose different combinations of methods to be available in UPAS.

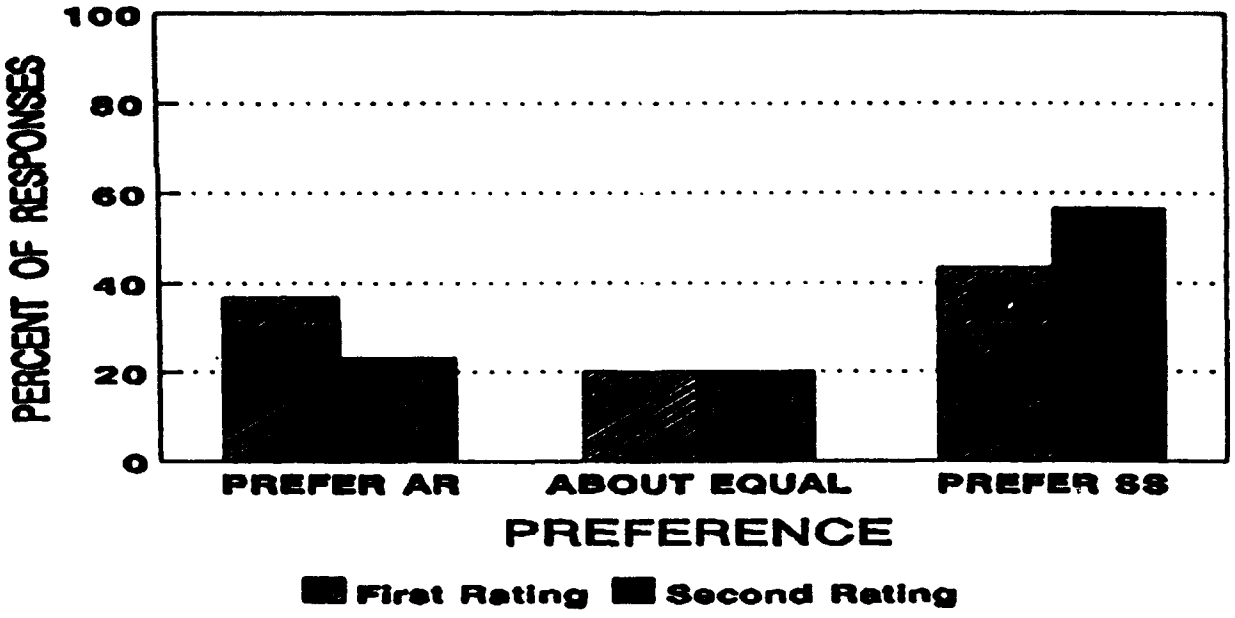


Figure 9. Change in preference for SIMNET exercise replay methods between ratings before and after seeing items about method characteristics. The category "prefer" also includes responses of "much prefer" and "slightly prefer".

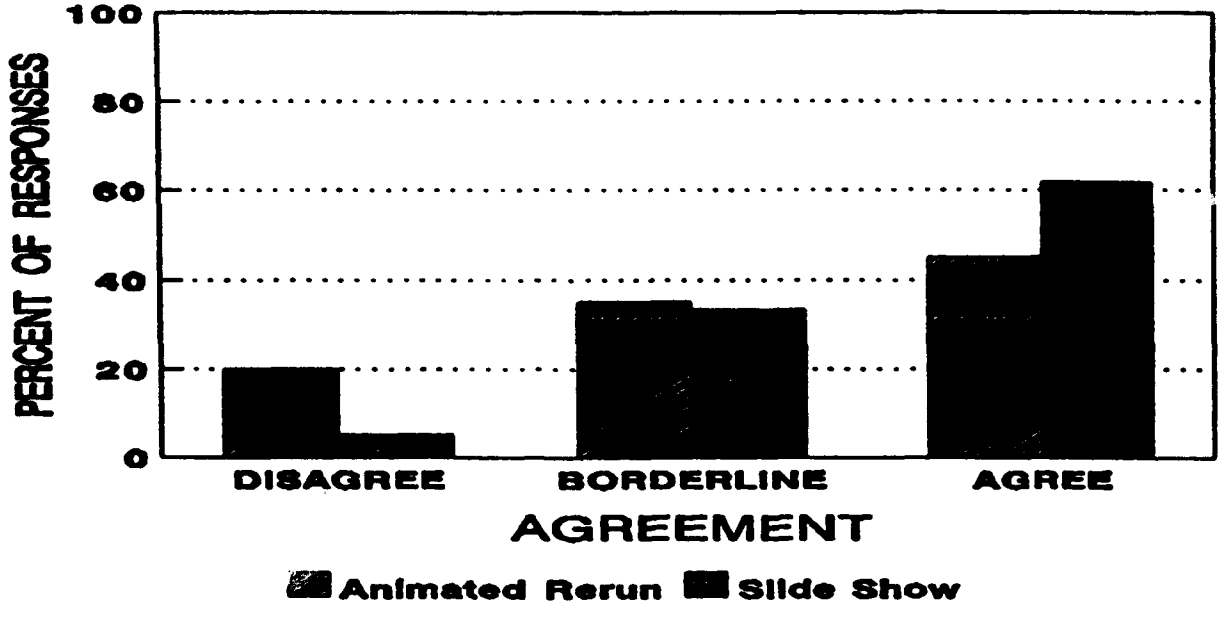


Figure 10. Agreement with statements that SIMNET exercise replay methods help to improve the effectiveness of an AAR. The category "agree" includes the response "strongly agree", and the category "disagree" includes "strongly disagree".

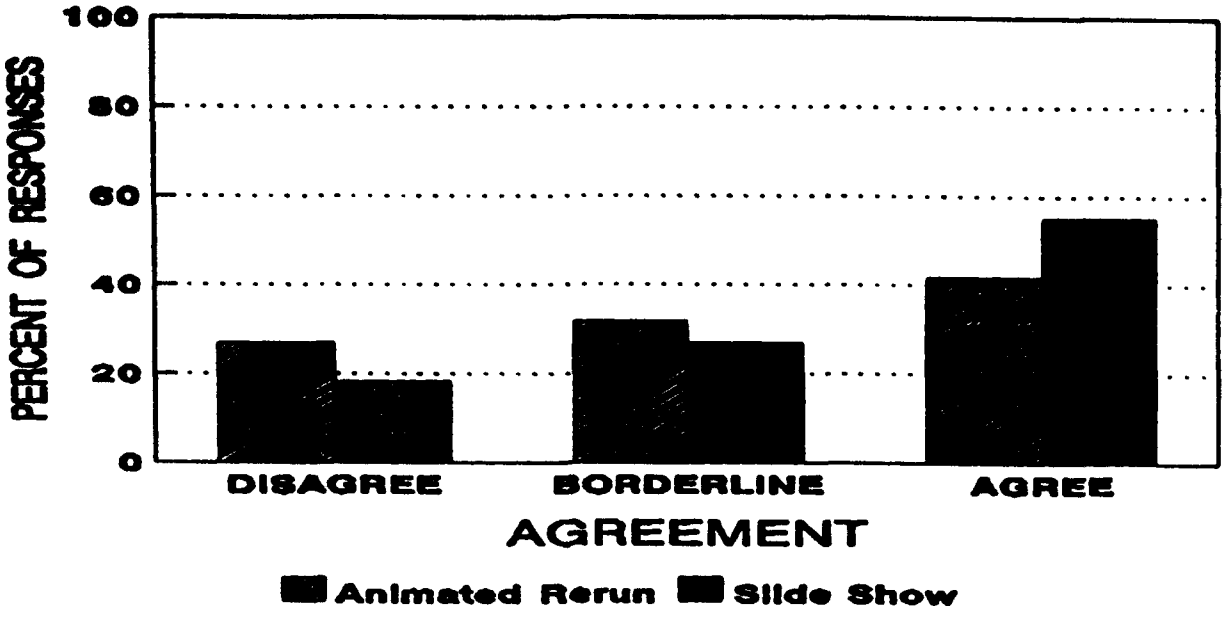


Figure 11. Agreement with statements that SIMNET exercise replay methods help AAR participants to pinpoint mission execution problems. The category "agree" includes the response "strongly agree", and the category "disagree" includes "strongly disagree".

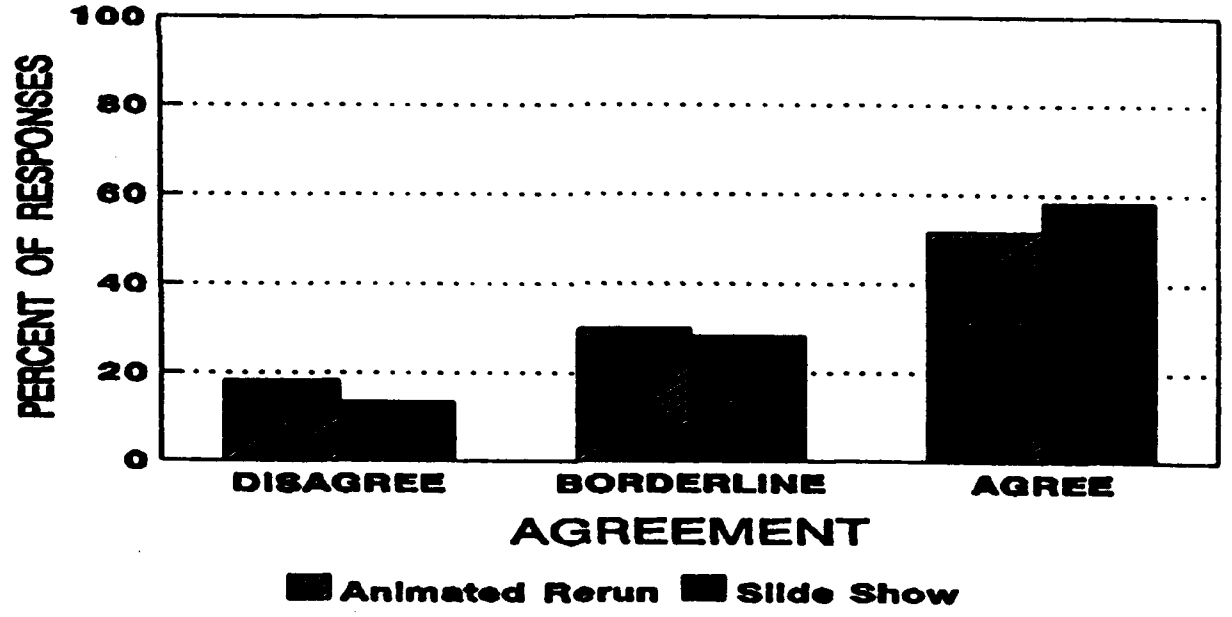


Figure 12. Agreement with statements that SIMNET exercise replay methods in AARs help the unit to improve mission performance. The category "agree" includes the response "strongly agree", and the category "disagree" includes "strongly disagree".

show replay was found for the statement--"Using this display would take too much time away from the AAR." Fifteen percent of the instructors concurred with this statement for the slide show method as compared to 38% who thought that the animated rerun would take too much time away from their AARs (see Figure 13). The instructors' average estimate was 11.57 minutes for the maximum amount of time that would be allotted for presenting these displays. Also, significant differences supporting the slide show display were found on the question dealing with the system's operational time (see Figure 14). Two-thirds of the instructors liked the fact that the slide show's operational time was based on the number of events that the AAR leader chose to show. Only 23% of them liked the temporal aspects associated with operating the animated rerun.

A significant difference favoring the slide show method was found for the questions on operational procedures. Two-thirds of the instructors liked the procedures required to operate the slide show method while 57% disliked the animated rerun's procedures (see Figure 15). The instructors thus felt that the slide show method was easier to use than the animated rerun method.

These data also provided some insights into the instructors' desires for having both replay methods available to use in SIMNET AARs. The instructors felt that both methods would be helpful to their students. As previously noted, fifty-eight percent and 51% of the instructors 51 indicated that the slide show and animated replay methods, respectively, would help students to improve their mission performance. System control was an important criterion in their willingness to utilize the different replay methods. Ninety and 84% of them liked the sense of control associated with the slide show and animated rerun methods, respectively (see Figure 16). These percentages were substantially higher than those found for the other features.

Divergent ratings were found regarding the instructional effectiveness of these replays. For example, the AOBC NCOs leaned toward strongly agreeing (mean = 4.18) with the statement that the slide show would help the unit find ways to improve their mission performance. The AOBC officers tended to neither agree nor disagree with this statement (mean = 2.78). The AOBC officers who conducted the AARs for the AOBC classes might then have had more realistic expectations about the effects of these displays upon students' performance than did the AOBC NCOs.

The reader is referred to Appendix E for a further description of the results associated with instructional effectiveness and the statement about taking too much time away from the AARs. Technical descriptions of the other results are presented in Appendix F.

Supporting Display Results. These results showed that the instructors favored the Battle Snapshot and graphic materials

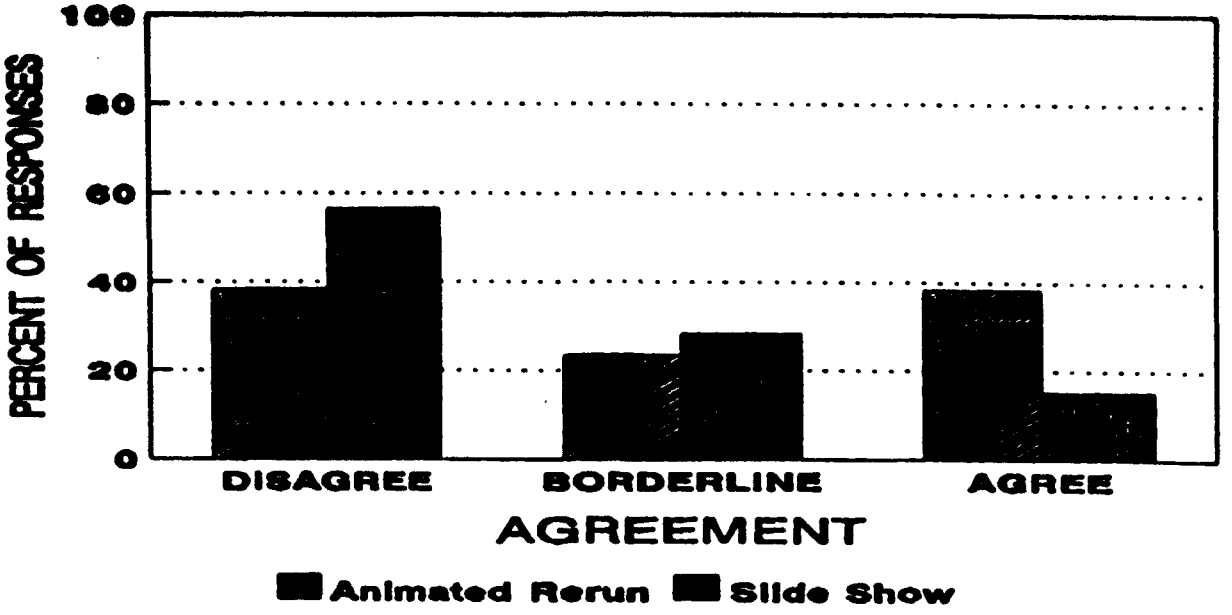


Figure 13. Agreement with statements that SIMNET exercise replay methods would take too much time away from the AAR discussion. The category "agree" includes the response "strongly agree", and the category "disagree" includes "strongly disagree".

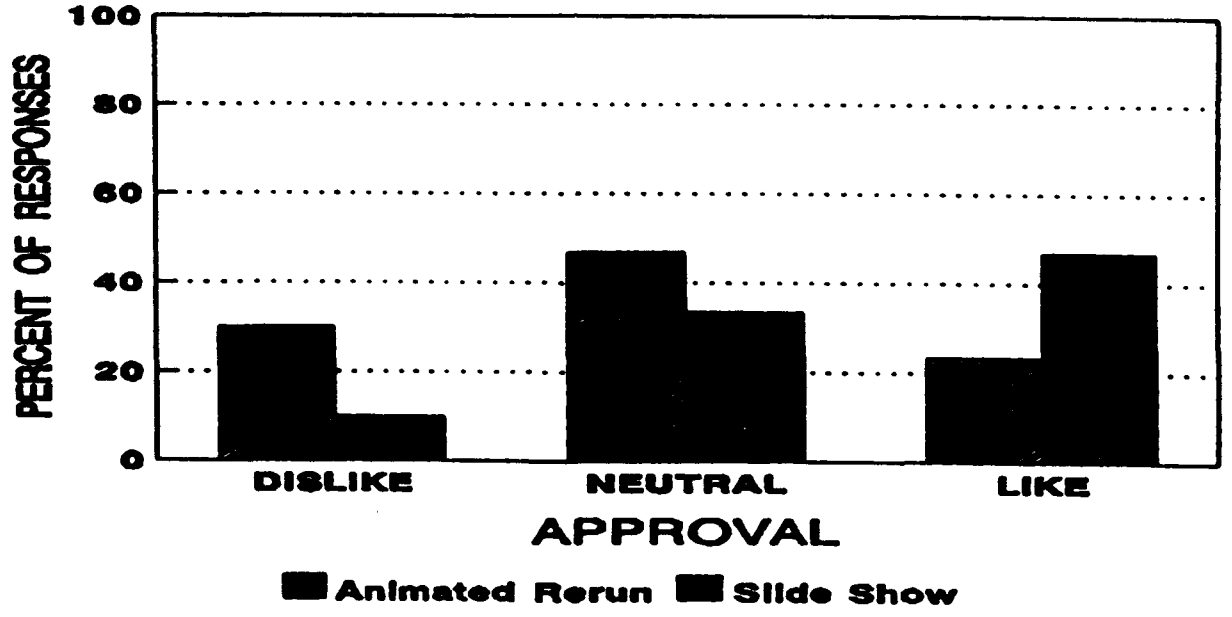


Figure 14. Approval of how the presentation time varies for each of the SIMNET exercise replay methods. The categories "like" and "dislike" include the response degrees "strongly like/dislike" and "slightly like/dislike".

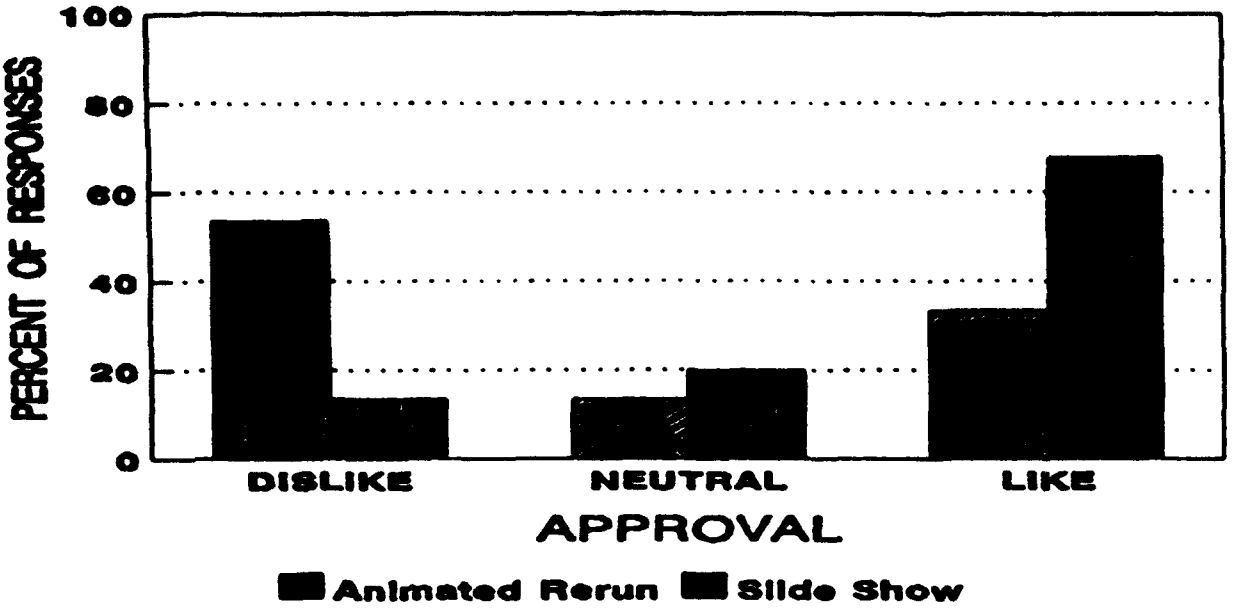


Figure 15. Approval of how the presentation operates for each of the SIMNET exercise replay methods. The categories "like" and "dislike" include the response degrees "strongly like/dislike" and "slightly like/dislike".

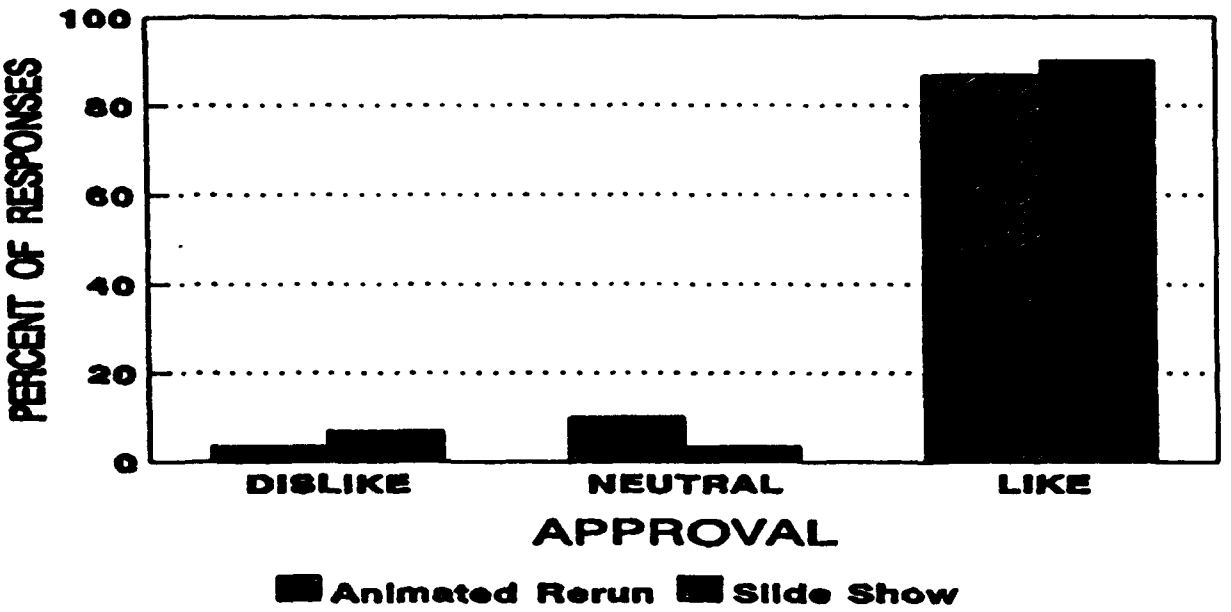


Figure 16. Approval of the type of control provided by each of the SIMNET exercise replay methods. The categories "like" and "dislike" include the response degrees "strongly like/dislike" and "slightly like/dislike".

over the Battle Flow and Exercise Timeline. Over 55% of them wanted to keep the Battle Snapshot with/without line-of-sight and the graphs as part of the UPAS package. Fewer than 45% wanted to keep the Exercise Timeline and Battle Flow (see Figure 17). This preference pattern was also manifested in their mean ranking data with the Battle Snapshot and Exercise Timeline having the highest rankings (see Figure 18).

The relatively low scores for the Battle Flow and Exercise Timeline displays were a function of problems that the officers had with these displays. Figures 19 and 20 show that officers did not feel that these displays were needed while the NCOs did. The Battle Flow and Exercise Timeline displays also had lower priority ranks by officers compared to NCOs. A similar result was manifested again in a significant interaction found in the instructors' "usefulness" responses. As indicated in Figure 21, the AOAC and AOBC officers considered the Exercise Timeline and Battle Flow to be slightly useful while they considered the other displays to be fairly useful. The NCOs felt that all the displays were fairly useful. Perhaps, these divergent findings most likely reflected differences in training requirements for armor NCOs and officers with the Exercise Timeline and Battle Flow not being very useful for the officers.

An interaction was also found in the "usefulness responses" between missions and AAR displays (see figure 22). The Battle Snapshot with/without LOS and graphs were rated fairly useful for the company, offense, and defense missions but only somewhat useful for the TRMs; the Exercise Timeline and Battle Flow were rated as somewhat to slightly useful for the different force-on-force missions, but only somewhat useful for the TRMs. The Exercise Timeline was rated fairly to somewhat useful for the TRMs, but only somewhat to slightly useful for the different force-on-force missions. The Battle Flow was rated as somewhat to slightly useful for all missions. These findings have demonstrated the relative usefulness of the Battle Snapshot and graphic displays for most but not all SIMNET missions.

The supporting display data results have thus indicated that all displays would be used in ANCOC. The Battle Snapshot and graphic display would be more likely used in AOBC and AOAC than would the Battle Flow and Exercise Timeline displays. (See Appendix G for a more technical description of these results.)

Suggestions for Improvements: The instructors made 337 comments. Eighty percent of these comments (271) dealt with needed enhancements to the system. And, 84% of these requests dealt with enhancements to:

1. Vehicle Icons--84 comments for this category with the need for vehicle bumper numbers mentioned 39 times.

2. System Control Functions--75 comments for

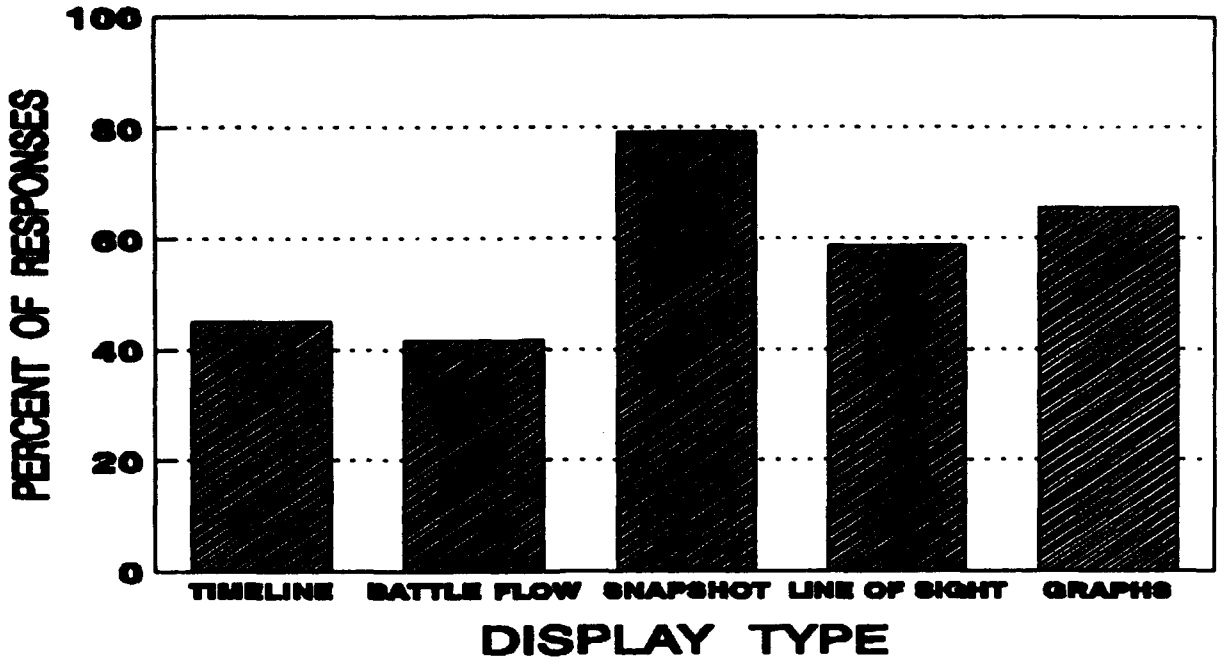


Figure 17. Percentage of responses that favor having each display in UPAS for use in AARs. Responses in the "yes" category are shown. Other response categories were "neutral" and "no".

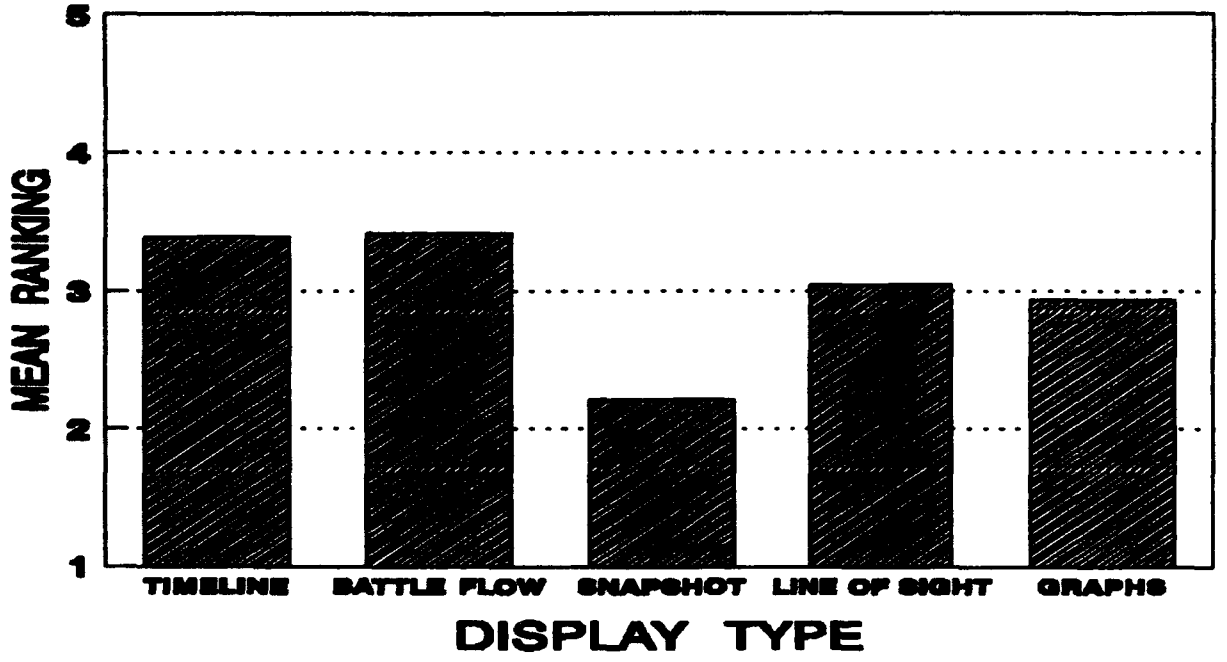


Figure 18. Priority ranking for keeping each display in UPAS to use in AARs. Lower rank value corresponds to higher priority.

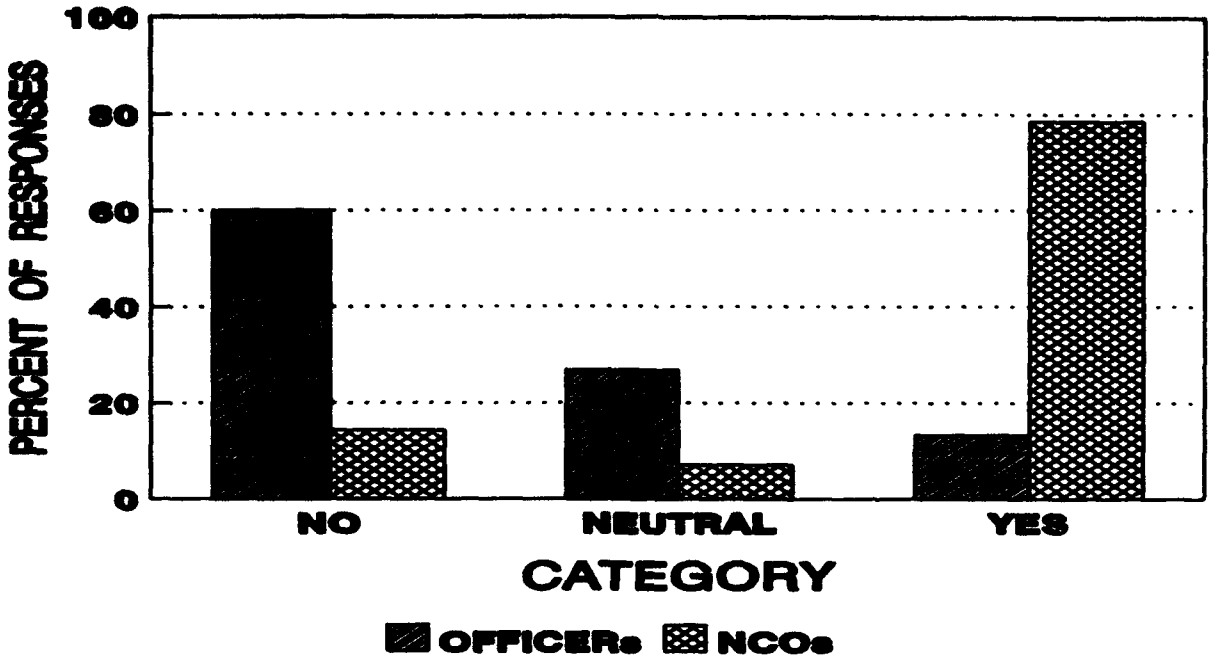


Figure 19. Differences in judgements between instructor groups about including the Timeline display in UPAS for use in AARs.

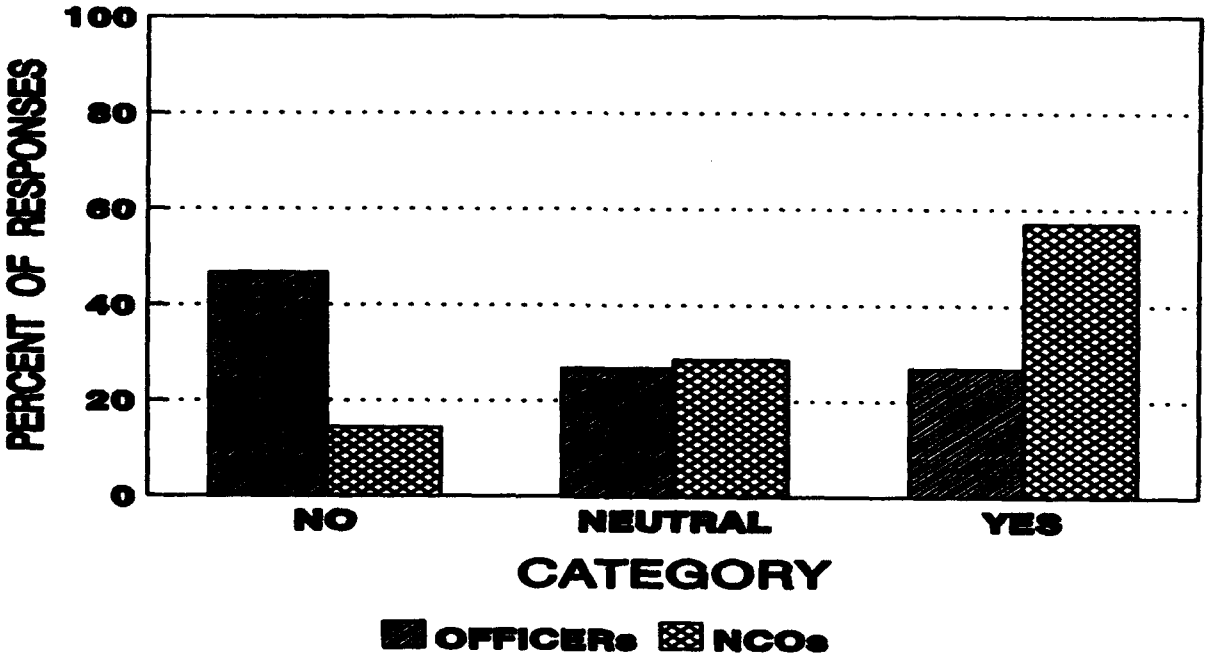


Figure 20. Differences in judgements between instructor groups about including the Battle Flow display in UPAS for use in AARs.

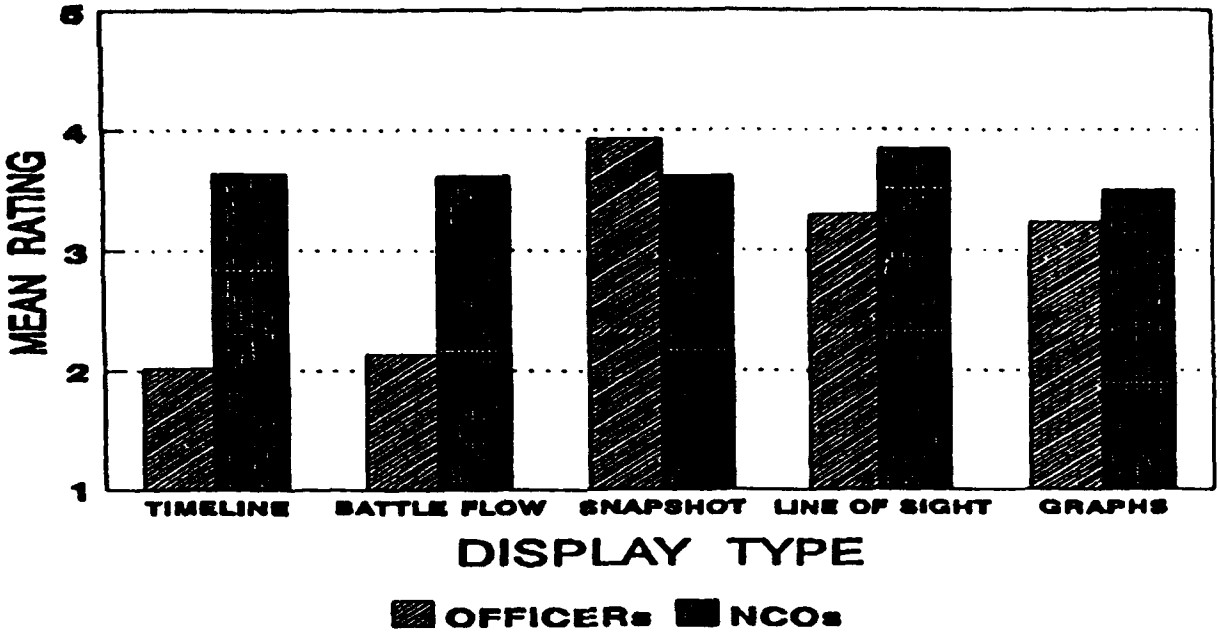


Figure 21. Usefulness of UPAS displays in AARs as rated by instructor groups. The responses on the 5-point rating scale were: "not useful" (1), "slightly useful" (2), "somewhat useful" (3), "fairly useful" (4), and "very useful" (5).

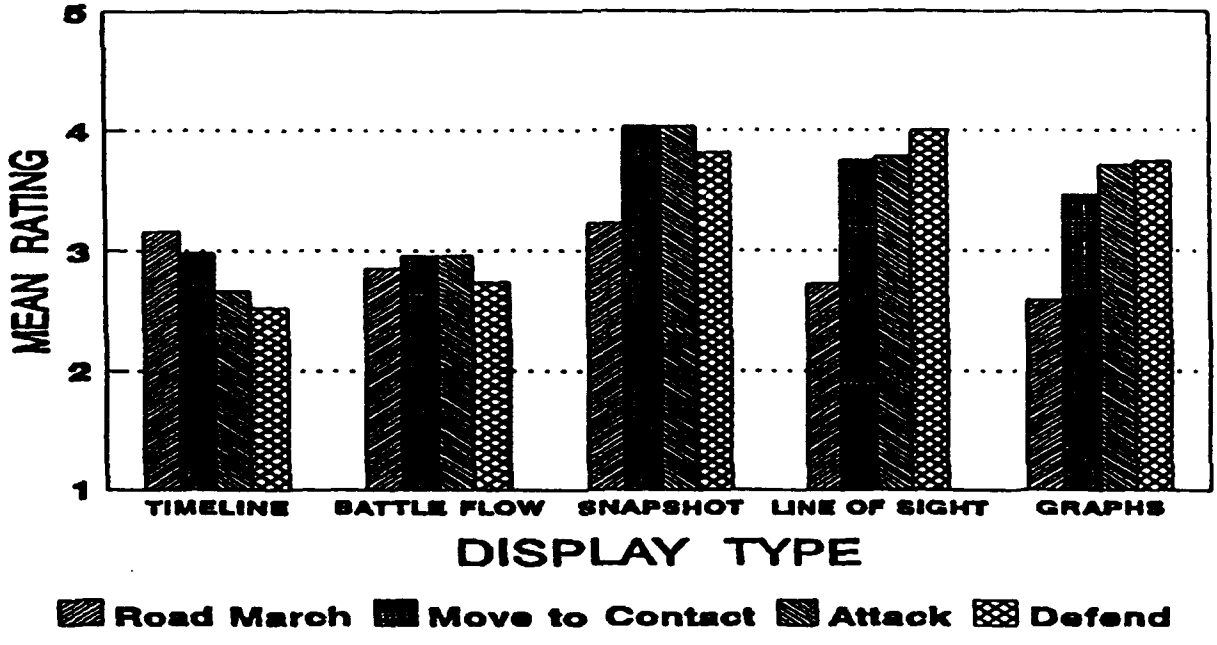


Figure 22. Usefulness of UPAS displays rated separately for four types of missions. Responses on the 5-point scale were: "not useful" (1), "slightly useful" (2), "somewhat useful" (3), "fairly useful" (4), and "very useful" (5).

this category with a need for a "fast-forward: mechanism mentioned 31 times. Also mentioned by a few instructors was the need for a mouse.

3. Map Features--68 comments for this category with the need for contour lines and better line-of-sight mentioned 14 and 10 times, respectively.

Ten comments also dealt with the need for more graphs.

These comments on suggested improvements indicated that UPAS' users want to be able to identify vehicles and quickly move through the action. By having a system that accomplished these two goals, UPAS would be easier for these trainers to use. That is, the users would be better able to identify specific vehicles for discussion purposes. The users would also waste less time in searching for the relevant screen(s) to view. Further improvements are thus needed to make UPAS easier to use and to help instructors meet certain training considerations (e.g., temporal constraints).

The instructors' comments were also broken down by questionnaire and display type. Over two-thirds of these comments (183) related to Questionnaire A dealing with the Plan View Display. Of the remaining 88 comments, 46 of them dealt with specific improvements to the animated replay method and 21 to the slide show method. Further evidence has thus been found for the previously discussed arguments that instructors found the slide show method more to their liking than the animated rerun method. And, that training considerations had a pronounced impact on their preference data.

The AOAC instructors, during the interview session, re-emphasized the need for new vehicle icons, system control features, map features, and graphs. They also emphasized the need for reply of communications within their units. Combining the UPAS with a communication replay was also mentioned by the instructors.

Summary of the Formal Assessment Data

The following implications can be drawn from these data:

1. Both the animated rerun and slide show displays would be utilized by these respondents with the slide show display being the preferred method;

2. The Battle Snapshot with/without LOS and graphic displays were the preferred supporting materials;

3. Training considerations had a pronounced impact on the instructors' preferences for the slide show, Battle Snapshot, and graphic displays;

4. The UPAS package, especially the Plan View Display, still needs some improvements with regards to its vehicle icons, control functions, and map features.

Validity Check Stage

Questions, however, remain about the validity of the formal assessment findings. It could be, for example, that the instructors' judgments about the UPAS displays reflected on the questions being asked rather than on any training needs. Hence, a validity check was needed of the questionnaires' data for actual SIMNET training situations.

Method

Participants. Five AOBC instructors participated in this phase. These instructors had participated in the formal assessment stage of this research effort.

Materials. The validity check's materials consisted of a structured think-aloud protocol and corresponding response sheets. (These materials can be found in Appendix K). A "think-aloud" procedure involved the instructors in commenting on their feelings, needs, or cognitive behaviors as they completed the task. Such protocols have been widely used for assessing subjects' responses to naturally occurring events or situations (see Erickson & Simon, 1984 for a detailed discussion of this procedure). Recently, variations of this technique have been used to measure research participants' acceptance of computer-based instructional materials (Shlechter, Burnside, & Thomas, 1988).

Procedure. Comments from these AOBC instructors were gathered as their students completing the following missions: (a) TRMs, (b) FOF, (c) CL ATK, and (d) CL DEF. Before starting these exercises, these instructors were carefully briefed by three trained ARI personnel on the procedures associated with this task. These instructors were shown, as a "memory refresher," pictures of the different displays.

The ARI personnel wrote down the instructors' comments on the response sheet. They were instructed to prompt, as unobtrusively as possible, the instructors to comment when comments were not forthcoming. They were also instructed to refrain from asking questions or making comments until the particular exercise was finished.

Two cycles of the AOBC course were sampled, which involved 24.09 hours worth of MWSTC time. This time was broken down as: (a) 9.4 hrs for TRMs, (b) 5.36 hrs for OFF, (c) 4.33 hrs for DEF, (d) 1.35 hrs for CL ATK, and (e) 2.25 hrs for CL DEF. Except for one instructor, each cycle involved a different set of AOBC instructors.

Scoring Procedure. Two independent judges conducted content analyses of these data. Separate analyses were conducted for each exercise and for the aggregate data. The categories for these analyses were the different types of requested display, including a category for non-UPAS or "other" displays, and the following activities: (a) troop leading procedures, (b) formation, (c) movement, (d) exercise of fires, (e) action on contact; (f) reports to commander, and (g) fratricides.

Very little difference was found in the content analyses performed by these two judges. Areas of disagreement were discussed resulting in a consensus decision on the correct classification of each response.

Results and Discussion

These instructors made 220 requests for an AAR aid (see Figure 23). The majority of these requests were for the Battle Snapshot displays with the other displays being requested between 5 and fourteen percent of the total. These instructors also wanted to have these different AAR displays available for use with the Battle Snapshot displays chosen as the most desirable. This finding paralleled the formal assessment findings on the usability of these different displays.

As also shown in Figure 23, these instructors made a sizeable number of requests for having other types of displays available for their AARs. Seventy-five percent of these "other" requests were for displays/graphs which would portray their students' communications (Commo) and fire fight performances. Again, the instructors professed a need for enhancing the UPAS package by developing communication and fire fight displays.

The think-aloud data also indicated a possible relationship between the instructors' comments and training requirements. Actions on contact and communication represented 40% and 20% of the SIMNET activities, respectively. Also, some variability was found in the instructors' requests as a function of either the training exercise or task. While the Battle Snapshot displays were preferred for most exercises and activities, the Exercise Timeline was the preferred display for troop leading procedures; the percentage of Battle Flow requests was much greater for the TRMs than for the other exercises.

The Validity Check results have provided the R & D team with further confidence that the UPAS displays, especially the Battle Snapshot displays, would be utilized by institutional trainers. These data have also indicated that enhancements were needed to the UPAS package.

Generalizability Check Stage

Questions, however, remain about the generalizability of the formal assessment findings. That is, the previously stated

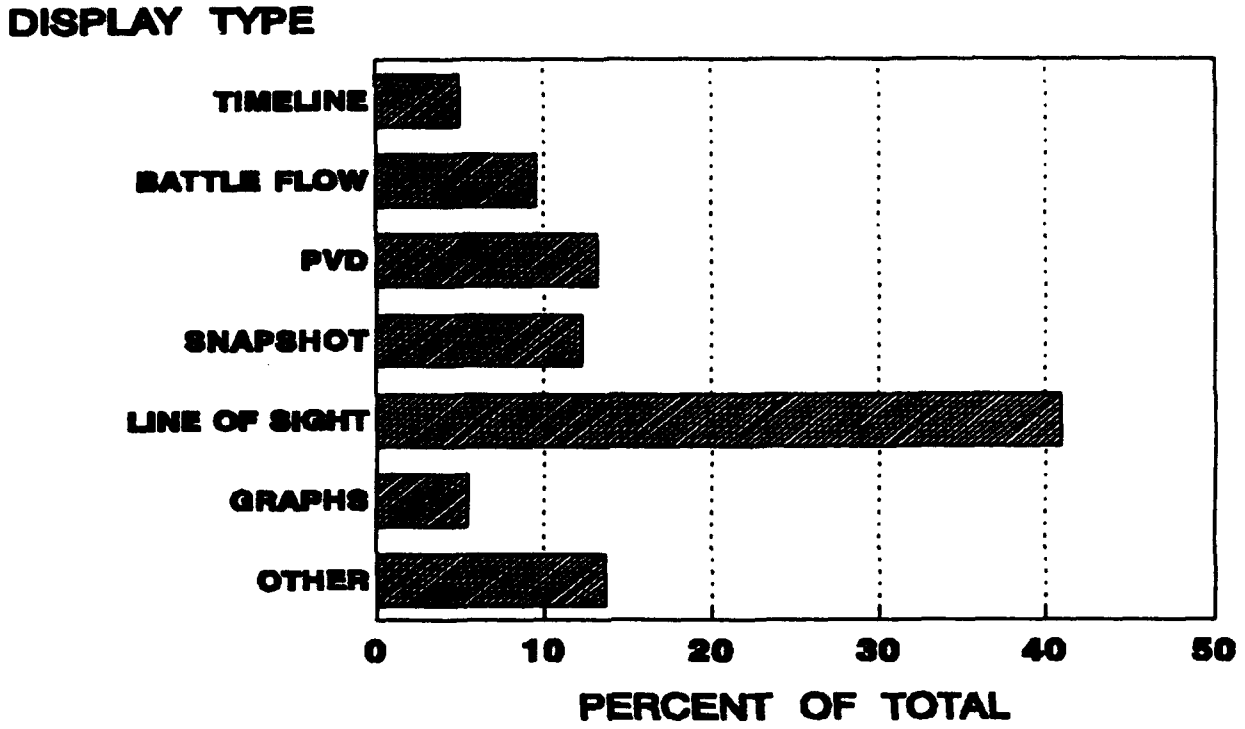


Figure 23. Relative requests for UPAS displays by AOB instructors that were observed during ongoing SIMNET exercises. The total number of requests was 220.

judgements about UPAS may not have reflected those of unit trainers. Hence, the UPAS' displays may not be appropriate for such training needs[4].

Method

The participants were seven company-grade officers from an armor battalion, who annually used the MWSTC as part of their NTC preparation. They were presented a modified version of the same displays which were presented to the formal assessment participants. After viewing these materials, they then completed the preference questionnaire which was used in the preliminary evaluation[5]. Time constraints were the reasons for these departures from the formal assessment stage. Time constraints also prevented any in-depth interviews.

---

[4] The data for this stage were collected prior to the data collection for the formal assessment data.

[5] The Battle Snapshot for this stage did include the line-of-sight vectors.

Findings

As shown in Figure 24, these participants felt that the Battle Snapshot and Battle Flow were fairly to extremely useful, while the Exercise Timeline, Plan View Display, and graphic displays were slightly to fairly useful. These data have thus provided the research and development team with further confidence regarding the utility of these UPAS displays. That is, the UPAS displays--especially the Battle Snapshot and Battle Flow displays--would seemingly be utilized by unit trainers.

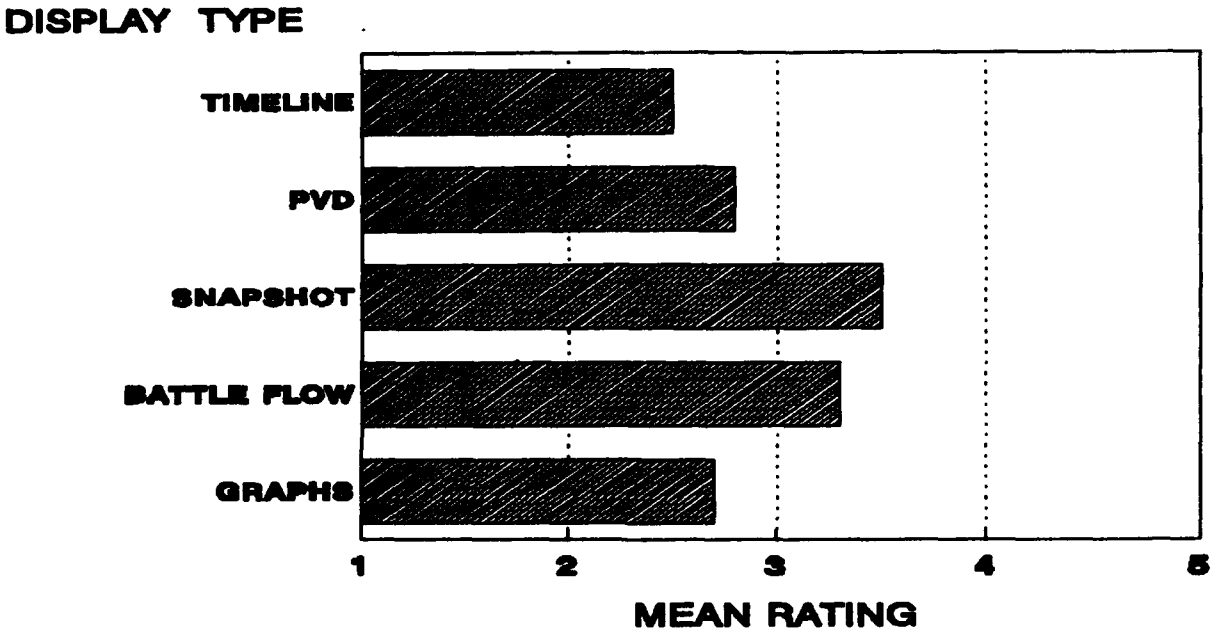


Figure 24. Usefulness of UPAS AAR aids rated by unit trainers. Responses on the 5-point rating scale were: "not useful" (1), "slightly useful" (2), "fairly useful" (3), "extremely useful" (4), and "indispensable" (5).

General Discussion

This research effort has shown that training personnel utilizing the MWSTC would like to have the different UPAS displays available, including both versions of the Plan View Display, for use in their AARs. Also, the slide show method was found to be preferable to the animated rerun method. The Battle Snapshot displays were likely to be used more than the Exercise Timeline and Battle Flow. These cited preference findings tended to occur across the different types of participants and methodologies.

Insights into the reasons for these cited preference data were also found. As discussed, the preference for the slide show method was a function of the display's apparent ability to accommodate certain training considerations and to be easy to use. Meeting training needs also seemed to be a determining

factor in the preference for the Battle Snapshot displays. Further evidence has thus been provided for the argument that instructional personnel must view any instructional innovation as meeting their needs and being easy to use; otherwise, they would not readily accept the innovation.

This research effort has also shown the value of R & D personnel working with the intended users. As discussed, certain refinements were made to the existing materials based on the feedback provided by these practitioners. The Battle Snapshot, for example, was modified to include line-of-sight vectors. The participants' comments were also instrumental in the development of additional UPAS materials, e.g., the recent development of a fire fight display. This partnership between developers and practitioners has helped make UPAS as user-friendly and responsive to users' needs as possible. Correspondingly then, UPAS has been refined through these users' feedback to improve support for their AARs.

This partnership has also helped determine the requirements for future UPAS-type systems, e.g., the Automated Training Analysis and Feedback System (ATAFS). The ATAFS will provide a prototype for the feedback system in the Close Combat Tactical Trainer. ATAFS is expected to include many capabilities requested by users that could not be developed within limitations of the UPAS hardware platform. The ATAFS will be a multi-tasking system, allowing simultaneous data recording and monitoring of the exercise on the PVD map. Recording and playback of communications will be synchronized with recording and playback of network data. The system also will include programmed expert assistance that will prompt, guide, and help manage the preparation of AAR displays. Many of the displays will be processed in real-time during the course of the exercise, so they will be available immediately for AAR use. And, this system is expected to include a mouse which would allow the users to select the vehicles for which bumper numbers are to be displayed.

This partnership, however, must not end with this research effort. Future efforts between the R & D team and the instructional personnel are needed to determine the most appropriate strategies for implementing UPAS into the different training programs which use the MWSTC. Implementation trials of the UPAS package are also needed. This later effort will involve collecting data on the: (a) instructors' use of UPAS displays for an actual AAR, (b) instructors' problems with using these materials, and (c) the students' attitudes toward the UPAS package.

A tentative framework for involving instructional personnel in the research and development process has emerged from the Army Research Institute's development and refinements of UPAS. This framework consists of employing users' input to help: (a) design the prototypes, (b) determine their instructional suitability, (c) refine the prototypes, (d) design the implementation process,

(e) test the implementation process, (f) make any final improvements to the package or implementation process, and (g) help guide the research and development of the next generation of this automated performance measurement system.

Many issues, however, need to be resolved concerning this framework. One such issue is the role of perspective users during the early stages of the research and development process. As stated, both the Battle Flow and Battle Scorecard were designed based on user comments. However, these two displays were not among those highly valued by most of the research participants. Perhaps, the research and development team should have, initially, obtained feedback from trainers about the unit performance measures that needed to be measured. And then, the design displays would have been created to meet these needs. The research and development team did not purposefully exclude potential users from early stages of the design process. They were excluded because their chain-of-command was not in favor of an automated performance measurement system for SIMNET.

It may have been relatively unproductive to have users evaluate prototype displays as was done in the Shlechter, et al. (1992) research effort. For one thing, trainers have too little experience with data displays to create such displays. Secondly, instructors may only be able to judge the value of an instructional feature based on an actual training condition. This point was duly noted by an armor school instructor. Another instructor claimed that he had to play with the UPAS device before evaluating it. Users seemed to need concrete examples of products stimulating more meaningful comments regarding the product's instructional value.

However, comments on prototype displays may not be totally worthless. Shlechter et al.'s (1992) data did indicate possible problems with the developed Plan View Display. Their subjects also claimed that this display needed to be enlarged. These subjects also indicated the need for the Plan View Display, Battle Flow, and Battle Snapshot displays to include terrain features. Perhaps then, users' comments on prototype displays can be viewed as needed formative evaluation data; while summative data from users must wait until the product is more fully developed.

Another unresolved issue involves the number of users who should be included in the research and development process. Sampling too many instructors may complicate this process as each may have his/her own instructional concerns. Not sampling enough instructors, however, may produce an instructional product which is not widely used.

This report has both theoretical and practical ramifications for instructional developers, practitioners and researchers. As discussed, a seemingly usable set of automated feedback materials for SIMNET exercises has been developed. Also discussed was the

partnership among instructional developers, practitioners and researchers in this R & D process. The authors believe that a productive partnership among these different professionals would reverse the previously discussed historical trend of use and then abandonment of new instructional innovations.

Conclusions

This research supports the following conclusions:

1. Training personnel at the MWSTC would likely use the UPAS materials. However, as shown in the results section, the Armor School instructors did not show high approval of the UPAS displays for this stage of the system's development.

2. Training personnel at the MWSTC would more likely use the slide show Plan View Display method than the animated rerun method.

3. Training personnel at the MWSTC would most likely use the Battle Snapshot displays more than the Exercise Timeline and Battle Flow displays.

4. Further improvements are needed to make UPAS easier to use and to speed production of AAR materials.

References

Alluisi, E. A. (1991). The development of technology for collective training. Human Factors, 33, 343-362.

Babbitt, B. A., & Nystrom, C. O. (1989). Questionnaire construction manual (ARI Research Product 89-20). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A212 365)

Bessemer, D. W. (1991). Transfer of SIMNET training in the Armor Officer Basic Course (ARI Technical Report 920). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A233 198)

Brown, R. E., Pishel, R. E., & Southard, L. D. (1988, April). Simulation networking (SIMNET) preliminary training developments study (PTDS) (TRAC-WSMR-TEA-8-88). White Sands Missile Range, NM: U.S. Army TRADOC Analysis Command. (Available from the U.S. Army Armor School, ATTN: ATSB-DOTD-ORA, Fort Knox, KY 40121-5200)

Burnside, B. L. (1990). Assessing the capabilities of training simulations: A method and simulation networking (SIMNET) application (ARI Research Report 1565). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A226 354)

Cochran, W. O., & Cox, G. M. (1957). Experimental designs. New York: Wiley.

Cuban, L. (1986). Teachers and machines: The classroom use of technology since 1920. New York: Teachers College Press.

Defense Science Board. (1976, February). Summary report of the DSB Task Force on Training Technology. Washington, DC.

Defense Science Board. (1988, May). Report of the DSB Task Force on Computer Applications to Training and Warfare. Washington, DC.

Downs, C. W., Johnson, K. M., & Fallesen, J. J. (1987). Analysis of feedback in after action reviews (ARI Technical Report 745). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A188 336)

Easter, A. W., Kryway, J. T., Olson, W. R., Peters, S. M., Slemon, G. K., & Obermayer, R. W. (1986). Instructor support feature guidelines (AFHRL-TR-85-57(II)). Brooks Air Force Base, TX: U.S. Air Force Human Resources Laboratory.

Erickson, K. A., & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. Cambridge, MA: The MIT Press.

Garlinger, D. L. (1987). Feedback principles for command group training (ARI Research Report 1459). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A191 860)

Garvey, E. E., & Radgowski, T. (1988). SIMNET-D standing operating procedure (Report No. 6929). Cambridge, MA: BBN Systems and Technologies Corporation.

Gray, T., Roberts-Gray, C., & Gray, W. D. (1983). A guide to implementation of training products (ARI Research Report 1350). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A143 669)

Holding, D. H. (1965). Principles of Training. Oxford, England: Pergamon Press.

Kepple, G. (1973). Design and analysis: A researcher's handbook. Englewood Cliffs, NJ: Prentice-Hall.

Kristiansen, D. M. (1987). SIMCAT: A platoon-level battlefield simulation (ARI Research Report 1439). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A182 172)

Kulhavy, R. W., & Stock, W. A. (1989, February). Feedback in written instruction: The place of expectancy (Instructional Research Faculty Research Report No. 14). Tempe, AZ: Arizona State University.

McFann, H. H., Hiller, J. H., & McCluskey, M. R. (1990, June). Unit performance measurement: Concepts, techniques and examples. Paper presented at the Military Operations Research Society.

Meliza, L. L., Bessemer, D. W., Burnside, B. L., & Shlechter, T. M. (1992). Platoon-level after action review aids in the SIMNET unit performance assessment system (UPAS) (ARI Technical Report 956). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A254 909)

Meliza, L. L., Tan, S. C., White, S., Gross, W., & McMeel, K. (1992). SIMNET unit performance assessment system (UPAS) user's guide (ARI Research Product 92-02). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A250 699)

Norusis, M., & SPSS, Inc. (1990a). SPSS/PC+ statistics 4.0 for the IBM PC/XT/AT and PS/2. Chicago: SPSS.

Norusis, M., & SPSS, Inc. (1990b). SPSS/PC+ advanced statistics 4.0 for the IBM PC/XT/AT and PS/2. Chicago: SPSS.

Polzella, D. J., Hubbard, D. C., Brown, J. E., & McLean, H. C. (1987). Aircrew training devices: Utility and utilization of advanced instructional features (AFHRL-TR-87-21). Brooks Air Force Base, TX: U.S. Air Force Human Resources Laboratory.

Polzella, D. J., & Hubbard, D. C. (1986). Aircrew training devices: Utility and utilization of advanced instructional features (Phase III--Electronic Warfare Trainers) (AFHRL-TR-85-49). Brooks Air Force Base, TX: U.S. Air Force Human Resources Laboratory.

Pridemore, D. R., Webb, J. M., Haygood, R. C., Stock, W. A., & Kulhavy, R. W. (1990). Knowledge estimation and post-response feedback. Paper presented at the annual meeting of the American Psychological Association, Boston, MA.

Reidel, S. L. (1988). User acceptance and field implementation of decision support systems (ARI Research Report 1477). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A200 412)

Roberts-Gray, C. (1983). On closing the implementation gap: Symposium proceedings (ARI Research Report 1344). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A168 643)

Scott, T. D. (1983). Tactical engagement simulation after action review guidebook (ARI Research Product 83-13). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A143 014)

Shlechter, T. M. (1991). Promises, promises, promises: History and foundations of computer-based training. In T. M. Shlechter (Ed.), Problems and promises of computer-based training (pp. 1-20). Norwood, NJ: Ablex Publishing Corporation.

Shlechter, T. M., Bessemer, D. W., & Kolosh, K. P. (1991). Role-playing and the effectiveness of a high-technology training simulator (ARI Technical Report 938). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A224 913)

Shlechter, T. M., Burnside, B. L., & Thomas, D. A. (1987). Issues in developing and implementing computer-based instruction for military training (ARI Research Report 1451). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A189 479)

Shlechter, T. M., Meliza, L. W., Burnside, B. L., & Bessemer, D. W. (1992, April). Developing automated feedback materials for a training simulator: An interaction between users and researchers. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Thomas, R. N. (1987). Educational radio and television--Their development in advanced industrial societies. In R. N. Thomas and V. N. Koybayashi (Eds.), Educational technology--Its creation, development, and cross-cultural transfer (pp. 104-125). New York: Pergamon Press.

TEXCOM Combined Arms Test Center. (1990, August). Close combat tactical trainer (CCTT) force development testing and experimentation (FDTE) (Test Report No. FD 0200). Fort Hood, TX. (Available from Headquarters, Training and Doctrine Command, ATTN: ATIS-S, Fort Monroe, VA 23651)

Appendix A

Preference Questionnaire for the Preliminary Assessment Stage

Based on your experience, answer the different questions for each instructional feature. Answers these questions in the context of usability for AARs. Assume each feature is equally easy to use and fully operational. Be as concise as possible when answering the open-ended questions.

1. Battle Scorecard

The Battle Scorecard contains:
too little information/too much information/the right amount of information.

reasons:

The Battle Scorecard is: not useful/slightly useful/fairly useful/extremely useful/indispensable.

reasons:

Any enhancements needed for the Battle Scorecard:

2. Exercise Timeline

The Exercise Timeline contains: too little information/too much information/the right amount of information.

reasons:

The Exercise Timeline is:not useful/slightly useful/fairly useful/extremely useful/indispensable.

reasons:

Any enhancements needed for the Exercise Timeline:

3. PVD

The PVD contains: too little information/too much information/the right amount of information.

The PVD is: not useful/slightly useful/fairly useful/extremely useful/indispensable.

Any enhancements needed to improve the PVD:

4. Battle Snapshot

The Battle Snapshot contains: too little information/too much information/the right amount of information.

The Battle Snapshot is: not useful/slightly useful/fairly useful/extremely useful/indispensable.

Any enhancements needed to improve the Battle Snapshot:

5. Battle Flow

The Battle Flow contains: too little information/too much information/the right amount of information.

The Battle Flow is: not useful/slightly useful/fairly useful/extremely useful/indispensable.

Any enhancements needed to improve the Battle Flow:

6. Graphs

The Graphs contain: too little information/too much information/the right amount of information.

The Graphs are not useful/slightly useful/fairly useful/extremely useful/indispensable.

Any enhancements needed to improve the Graphs:

Miscellaneous Questions

1. How potentially useful is UPAS as a take-home package?
not useful/slightly useful/ fairly useful/extremely useful/indispensable

2. Any enhancements needed to make it a take-home package?

3. Any other enhancements needed for UPAS?

UPAS Questionnaire A:
Questions on the Plan View Display

Please check one answer to indicate whether you agree or disagree with each of the following statements about the specific exercise replay that you just saw. Also, please give reasons for your opinion if you can.

If this Plan View Display (PVD) replay can be displayed on a large screen clearly visible to all exercise players in a platoon-level after action review (AAR):

1. This replay would help the AAR participants to easily see the sequence of events that occurred during the unit's mission.

___Strongly Agree ___Agree ___Borderline ___Disagree ___Strongly Disagree

Reasons:

2. This replay would help the AAR participants pinpoint the unit's problems in executing the mission.

___Strongly Agree ___Agree ___Borderline ___Disagree ___Strongly Disagree

Reasons:

3. This replay in the AAR would help the unit find ways to improve their mission performance.

___Strongly Agree ___Agree ___Borderline ___Disagree ___Strongly Disagree

Reasons:

4. Using this replay would help the trainer to identify tasks and skills that need additional training.

___Strongly Agree ___Agree ___Borderline ___Disagree ___Strongly Disagree

Reasons:

5. Using this replay would help improve the effectiveness of an AAR.

___Strongly Agree ___Agree ___Borderline ___Disagree ___Strongly Disagree

Reasons:

6. This replay would take too much time away from the AAR discussion.

___Strongly Agree ___Agree ___Borderline ___Disagree ___Strongly Disagree

Reasons:

UPAS Questionnaire B:
Summary Questions on Displays

A. Method of Presentation

The replays of sample exercises were presented by two different methods:
(1) an **animated rerun** of vehicle movement and fires through time, and
(2) a **slide show** of vehicle locations and fires at specific time points.

For each question, check one answer that best reflects your opinion about these methods. Your opinions are important to help develop the UPAS to meet your needs. Please give reasons for your opinion if you can.

1. Which method(s) would you like to have for use in AARs?.

☐ animated rerun only ☐ slide show only ☐ both methods ☐ neither method

Reasons:

2. How would an AAR conducted with the aid of a slide show compare to what you usually do in an AAR? Using a slide show would be:

☐	☐	☐	☐	☐	☐	☐
much worse	worse	slightly worse	about equal	slightly better	better	much better

Reasons:

3. How would an AAR conducted with the aid of an animated rerun compare to what you usually do in an AAR? Using an animated rerun would be:

☐	☐	☐	☐	☐	☐	☐
much worse	worse	slightly worse	about equal	slightly better	better	much better

Reasons:

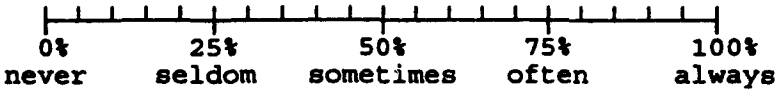
4. Would you prefer an animated rerun or a slide show as an AAR aid?

☐	☐	☐	☐	☐	☐	☐
much prefer (an animated rerun)	prefer	slightly prefer	about equal	slightly prefer (a slide show)	prefer	much prefer

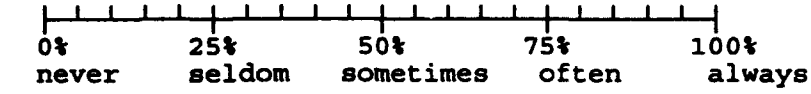
Reasons:

5. If only one method were available, how often would you use it? Mark an x on the scale to show your estimate of the percentage of AARs.

Slide Show:



Animated Rerun:



6. The slide show operates by pressing any key when you want to advance to the next display. Do you like or dislike this feature?

☐	☐	☐	☐	☐	☐	☐
strongly dislike	dislike	dislike slightly	neutral	like slightly	like	strongly like

7. The animated rerun operates by pressing function keys and entering numbers to change the terrain area or to jump ahead or back in time. Do you like or dislike this feature?

☐	☐	☐	☐	☐	☐	☐
strongly dislike	dislike	dislike slightly	neutral	like slightly	like	strongly like

8. The animated rerun takes less or more time based on the length of the exercises. Do you like or dislike this feature?

☐	☐	☐	☐	☐	☐	☐
strongly dislike	dislike	dislike slightly	neutral	like slightly	like	strongly like

9. The slide show takes more or less time based on the number of events the AAR leader chooses to show. Do you like or dislike this feature?

☐	☐	☐	☐	☐	☐	☐
strongly dislike	dislike	dislike slightly	neutral	like slightly	like	strongly like

10. The slide show lets the AAR leader try out and set up a fixed sequence of event displays before the AAR. Do you like or dislike this feature?

☐	☐	☐	☐	☐	☐	☐
strongly dislike	dislike	dislike slightly	neutral	like slightly	like	strongly like

11. The animated rerun lets the AAR leader choose when he wants to freeze the action, when to jump back and forth in time, and when and how to change the terrain area and scale. Do you like or dislike this feature?

☐	☐	☐	☐	☐	☐	☐
strongly dislike	dislike	dislike slightly	neutral	like slightly	like	strongly like

12. The slide show focuses on events the AAR leader picks, discarding unit actions considered unimportant. Do you like or dislike this feature?

☐	☐	☐	☐	☐	☐	☐
strongly dislike	dislike	dislike slightly	neutral	like slightly	like	strongly like

13. The animated rerun shows the unedited continuity of unit actions throughout the exercise. Do you like or dislike this feature?

☐	☐	☐	☐	☐	☐	☐
strongly dislike	dislike	dislike slightly	neutral	like slightly	like	strongly like

B. Time for After Action Reviews (AARs)

Please use your best judgment to estimate the following amounts of time. Mark an x on the scale to show your time estimate in terms of minutes. If your estimate is greater than the largest value shown, write the number of minutes in the blank provided.

If you complete a platoon-level exercise lasting 40 minutes and you plan to conduct an AAR afterward:

1. How many minutes do you **usually** give your soldiers for a break before starting the AAR?

0 5 10 15 20 25 30 or more_____ minutes

2. If preparing a PVD replay takes extra time, how much would you be willing to extend the break beyond the usual time (estimated in #1)?

0 5 10 15 20 25 30 or more_____ minutes

3. What is the **maximum** number of minutes you would ever give soldiers for a break before starting an AAR?

0 5 10 15 20 25 30 or more_____ minutes

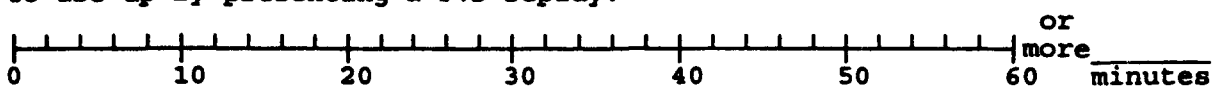
4. How much would you be willing to extend a break beyond the maximum (estimated in #3), if that is the only way you can get a PVD replay?

0 5 10 15 20 25 30 or more_____ minutes

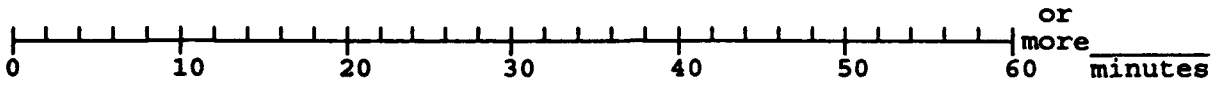
5. How many minutes do you **usually** spend with a platoon in a typical AAR after an exercise lasting 40 minutes? !!NOTE THE CHANGE OF SCALE!!

0 10 20 30 40 50 60 or more_____ minutes

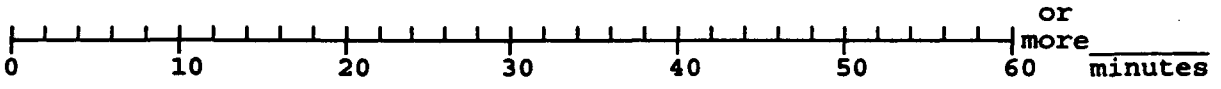
6. How many minutes of the usual AAR time (estimated in #5) would you want to use up by presenting a PVD replay.



7. What is the maximum number of minutes you would ever spend with a platoon in an AAR going over an exercise lasting 40 minutes?



8. How many minutes of the maximum AAR time (estimated in #7) would you be willing to use up by presenting a PVD replay?



C. After careful examination of different features and time factors, how strongly would you prefer an animated rerun or a slide show as an AAR aid?

☐ much prefer (an animated rerun) ☐ prefer ☐ slightly prefer ☐ about equal ☐ slightly prefer (a slide show) ☐ prefer ☐ much prefer

Reasons:

D. What suggestions do you have for improving the slide show method and/or the animated replay method?

UPAS Questionnaire C:
Questions on Different AAR Aids

A. Other UPAS Displays: Along with the Plan View Display (PVD), which other displays would you want to have for use in AARs? Also, please rank order your preference for **keeping** the displays in UPAS. Give your first priority rank 1, and your last priority rank 5.

Display	(Check one answer in each row) Yes	Neutral	No	Rank Order
Timeline	_____	_____	_____	_____
Battle Flow	_____	_____	_____	_____
Battle Snapshot	_____	_____	_____	_____
Line-of-Sight Snapshot	_____	_____	_____	_____
Graphs	_____	_____	_____	_____

2. How useful is the **Timeline** display in an AAR for each type of mission?

Mission	not useful	slightly useful	somewhat useful	fairly useful	very useful
Tactical Road March	_____	_____	_____	_____	_____
Move/React to Contact	_____	_____	_____	_____	_____
Attack	_____	_____	_____	_____	_____
Defense	_____	_____	_____	_____	_____

3. How useful is the **Battle Flow** display in an AAR for each type of mission?

Mission	not useful	slightly useful	somewhat useful	fairly useful	very useful
Tactical Road March	_____	_____	_____	_____	_____
Move/React to Contact	_____	_____	_____	_____	_____
Attack	_____	_____	_____	_____	_____
Defense	_____	_____	_____	_____	_____

4. How useful is the **Battle Snapshot** display in an AAR for each type of mission?

Mission	not useful	slightly useful	somewhat useful	fairly useful	very useful
Tactical Road March	_____	_____	_____	_____	_____
Move/React to Contact	_____	_____	_____	_____	_____
Attack	_____	_____	_____	_____	_____
Defense	_____	_____	_____	_____	_____

5. How useful is the **Line-of-Sight Snapshot** display in an AAR for each type of mission?

Mission	not useful	slightly useful	somewhat useful	fairly useful	very useful
Tactical Road March	_____	_____	_____	_____	_____
Move/React to Contact	_____	_____	_____	_____	_____
Attack	_____	_____	_____	_____	_____
Defense	_____	_____	_____	_____	_____

6. How useful are the **Graph** displays in an AAR for each type of mission?

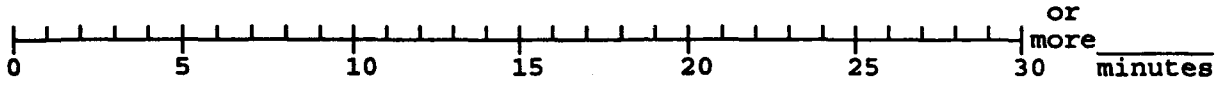
Mission	not useful	slightly useful	somewhat useful	fairly useful	very useful
Tactical Road March	_____	_____	_____	_____	_____
Move/React to Contact	_____	_____	_____	_____	_____
Attack	_____	_____	_____	_____	_____
Defense	_____	_____	_____	_____	_____

B. Time for After Action Reviews (AARs)

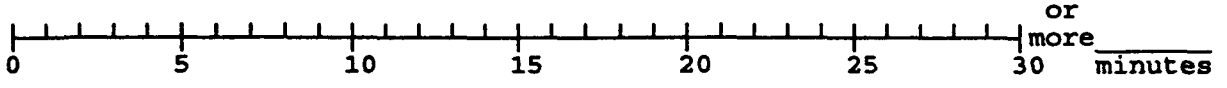
Please use your best judgment to estimate the following amounts of time. Mark an x on the scale to show your time estimate in terms of minutes. If your estimate is greater than the largest value shown, write the number of minutes in the blank provided.

If you complete a platoon-level exercise lasting 40 minutes and you plan to conduct an AAR afterward:

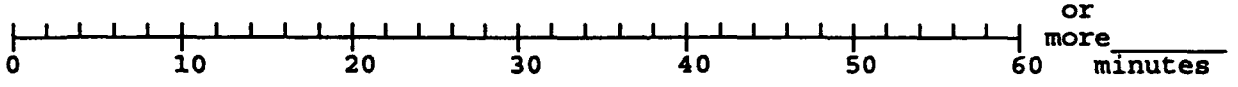
1. What is the **maximum** number of minutes you would ever give soldiers for a break before starting an AAR?



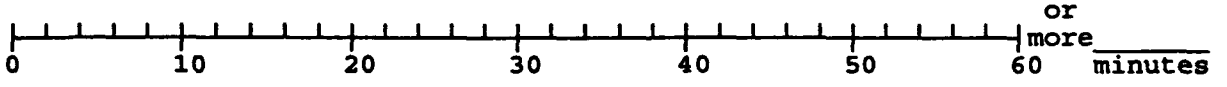
2. How much would you be willing to extend a break beyond the maximum (estimated in #2), if that is the only way you can get other types of displays in addition to a PVD replay?



3. What is the **maximum** number of minutes you would ever spend with a platoon in an AAR going over an exercise lasting 40 minutes?
!!NOTE THE CHANGE IN SCALE!!



4. How many minutes of the maximum AAR time (estimated in #2) would you be willing to use up by showing the PVD replay together with various types of other displays?



C. Do you have any final comments regarding UPAS?

Appendix E

Analysis of Responses to Questionnaire A

Conditions in Table E-1 show the Phase 3 research design that affected Questionnaire A. The research participants (respondents) saw four UPAS demonstrations of exercise replays based on four types of missions. After each replay, the respondents completed the questionnaire. The replays presented the exercise mission types in a constant order: (a) Tactical Road March (TRM), (b) Force-on-Force (FOF), (c) Attack (ATK), and (d) Defend (DEF). Combined with these missions, the demonstrations used both the animated rerun method and the slide show method for two replays. The methods appeared in different sequences for subgroups within each group of respondents. One subgroup saw the replays with the methods in Sequence 1 (animated rerun, slide show, slide show, animated rerun). The second subgroup saw the replays with the methods in Sequence 2 (slide show, animated rerun, animated rerun, slide show).

Table E-1
Research Design Used to Compare UPAS Replay Methods

	Order and Mission Type			
Instructor Group	1 TRM	2 FOF	3 ATK	4 DEF
AOAC Officers				
Sequence 1	AR	SS	SS	AR
Sequence 2	SS	AR	AR	SS
AOBC Officers				
Sequence 1	AR	SS	SS	AR
Sequence 2	SS	AR	AR	SS
ANCOC NCOs				
Sequence 1	AR	SS	SS	AR
Sequence 2	SS	AR	AR	SS
AOBC NCOs				
Sequence 1	AR	SS	SS	AR
Sequence 2	SS	AR	AR	SS

Note. TRM = Tactical Road March, FOF = Force-on-Force, ATK = Attack, DEF = Defense. AR = Animated Rerun Method, SS = Slide Show Method.

Questionnaire A contained six statements about the exercise replays presented to the participants. Participants expressed their agreement with the statements on a five-point scale ranging from "strongly disagree" to "strongly agree." Numerical scores 1-5 replaced these responses for statistical analysis. One participant did not respond to the fourth statement when he completed the questionnaire for the first time. For this nonresponse, the data analyses used the whole number nearest to an estimate of the missing value found by Yates' method (Cochran and Cox, 1957). The computation of the estimate used the data table only for that subject.

An initial multivariate analysis of variance (MANOVA) used the SPSS/PC+ procedure (Norusis & SPSS, 1990b) on the combined data for statements 1-6 (Item variable). The analysis included Groups and Sequence as between-subjects variables and Item and Order as repeated-measure variables. The analysis resulted in significant main effects for Group, $F(3, 21) = 7.91$, $p = .001$, and Item (Pillai's trace = .825), $F(5, 17) = 16.06$, $p = .000$. The Group by Item interaction also proved to be significant (Pillai's trace = 1.119), $F(15, 57) = 2.26$, $p = .014$. Such results show that the group differences vary by item. Separate analyses on the data for each item then used Group, Sequence, and Order variables. These analyses also provided estimates of orthogonal polynomial contrasts for the Order effects.

The layout in Table E-1 forms a counterbalanced crossover research design (Cochran and Cox, 1956). This design permits a precise statistical comparison between the animated rerun and slide show methods free of confounding with other factors manipulated in the design. In particular, this comparison is independent of overall differences between missions. This comparison is also independent of order differences that may result from repeated use of the same questionnaire. In the crossover design, the quadratic part of the Sequence by Order interaction is the same as the contrast between the animated rerun and slide show methods. The interaction contrast coefficients are 1, -1, -1, and 1 for Sequence 1, and -1, 1, 1, and -1 for Sequence 2. The pattern of coefficients coincides with the pattern of animated rerun and slide show methods in the two sequences. Therefore, an estimate of the quadratic interaction contrast is equal to the difference between slide show and animated rerun means.

Some confounding of Method by Mission interaction effects with Sequence by Order effects is possible in this crossover design. However, the Method by Mission effects must occur in a specific form to contaminate the comparison between animated rerun and slide show methods. This form appears when the differences between Sequences 1 and 2 (and methods) are (a) similar for the TRM and DEF missions, and (b) similar for the FOF and ATK missions. Furthermore, the differences for the TRM and DEF missions must be opposite to the differences for the FOF and ATK missions. Such a pattern of differences is very unlikely

given the nature of these missions. Without the animated rerun and slide show methods, no reasonable basis would exist for a difference in quadratic trend between Sequences 1 and 2. Therefore, the assumption is that the quadratic interaction effect can result only from a difference between methods.

The unweighted means for the animated rerun and slide show methods are shown in Table E-2 for all six questions. The unweighted means are averages of subgroup means based on differing numbers of participants. The MANOVA computes estimates and tests for effects using these means. Table E-3 presents the

Table E-2

Unweighted Means for UPAS Exercise Replay Methods

	Item					
Method	1	2	3	4	5	6
Animated Rerun	3.716	3.206	3.360	3.007	3.344	3.103
Slide Show	3.721	3.396	3.453	3.124	3.712	2.450

Note. Appendix C shows Questionnaire A, Items 1-6.

Table E-3

Significance Tests and Confidence Intervals on Contrasts Between Animated Rerun and Slide Show Methods of Presentation for Items on Questionnaire A

Item	Diff.	SE	t(22)	p	Lower-95% CL-Upper	
1	-0.005	0.147	-0.035	.972	-0.310	0.300
2	-0.300	0.174	-1.729	.098	-0.660	0.060
3	-0.093	0.137	-0.674	.507	-0.378	0.192
4	-0.117	0.091	-1.281	.214	-0.306	0.072
5	-0.368	0.141	-2.616	.016*	-0.659	-0.076
6	0.653	0.214	3.058	.006*	0.210	1.096

*$p < .05$.

MANOVA statistics for the *á priori* contrasts between these pairs of means.

The first four items concerned the helpfulness of the replays in after action reviews (AARs). For these items, the difference in means between methods is small and statistically insignificant. The fifth item stated that the replay would improve the effectiveness of the AAR. The significant difference in this case showed a larger mean, and thus stronger agreement, for the slide show method compared to the animated rerun method. For the sixth item, the statement asserted that the replay would take too much time. The significant method difference for the sixth item showed a larger mean and stronger agreement for the animated rerun method. With this negative feature, the smaller mean shows greater disagreement and is favorable to the slide show method.

Table E-4 lists item means for each group of participants. These means show that the average response to the replays varies widely among the groups for most Items. The pattern of differences is consistent for Items 1-3. The means are larger for AOAC Officers and AOBC NCOs, and are smaller for AOBC Officers and ANCOC NCOs. Other patterns appear for Items 4-6. Table E-5 gives the MANOVA test statistics for the between-subject effects for each item. Table E-6 presents statistics for the repeated-measure effects and interactions.

Table E-4

Unweighted Group Means for Questionnaire A Items

	Item					
Group	1	2	3	4	5	6
AOAC Officers	3.893	3.244	3.513	3.263	3.569	2.981
AOBC Officers	3.219	2.469	2.781	2.188	3.500	2.500
ANCOC NCOs	3.406	3.063	3.156	2.969	2.906	3.188
AOBC NCOs	4.354	4.408	4.177	3.844	4.135	2.438

Note. Appendix C shows Questionnaire A, Items 1-6.

Group differences are significant for all except Item 6. In Table E-5, every test of the Sequence by Group interaction is not significant. In Table E-6, tests of the Sequence by Group by Order interactions also are insignificant. The latter result permits Sequence by Order effects (including the differences between animated rerun and slide show methods) to be interpreted independently from the group differences. Conversely, group differences are independent of method differences. For Items 4
Table E-5

Between-Subjects Tests of Significance for Questionnaire A Items

Source of Variation	SS	df	MS	F	p
		Item 1			
SEQUENCE (S)	1.77	1	1.77	1.63	.216
GROUP (G)	20.55	3	6.85	6.30	.003*
S BY G	11.25	3	3.75	3.45	.034
WITHIN CELLS	23.90	22	1.09		
		Item 2			
SEQUENCE (S)	4.82	1	4.82	3.14	.090
GROUP (G)	39.24	3	13.08	8.53	.001*
S BY G	6.34	3	2.11	1.38	.276
WITHIN CELLS	33.74	22	1.53		
		Item 3			
SEQUENCE (S)	6.98	1	6.98	8.70	.007*
GROUP (G)	26.92	3	8.97	11.18	.000*
S BY G	2.39	3	0.80	0.99	.414
WITHIN CELLS	17.65	22	0.80		
		Item 4			
SEQUENCE (S)	10.73	1	10.73	7.48	.012
GROUP (G)	34.37	3	11.46	7.99	.001*
S BY G	5.71	3	1.90	1.33	.291
WITHIN CELLS	31.55	22	1.43		
		Item 5			
SEQUENCE (S)	.75	1	0.75	0.43	.519
GROUP (G)	22.52	3	7.51	4.30	.016*
S BY G	5.67	3	1.89	1.08	.378
WITHIN CELLS	38.43	22	1.75		
		Item 6			
SEQUENCE (S)	10.62	1	10.62	7.31	.013
GROUP (G)	11.44	3	3.81	2.62	.076
S BY G	5.92	3	1.97	1.36	.282
WITHIN CELLS	31.99	22	1.45		

*$p < (.01)df$.

Table E-6

Tests of Pillai's Trace Statistic for Repeated-Measure Effects

Source of Variation	Trace	F	df	Error df	p
		Item 1			
ORDER (O)	.03445	0.238	3.00	20.00	.869
SEQ. (S) BY (O)	.05840	0.413	3.00	20.00	.745
GROUP (G) BY (O)	.35400	0.981	9.00	66.00	.464
S BY G BY O	.27850	0.750	9.00	66.00	.662
		Item 2			
ORDER (O)	.07746	0.560	3.00	20.00	.648
SEQ. (S) BY (O)	.22457	1.931	3.00	20.00	.157
GROUP (G) BY (O)	.47004	1.362	9.00	66.00	.223
S BY G BY O	.35576	0.987	9.00	66.00	.459
		Item 3			
ORDER (O)	.14086	1.093	3.00	20.00	.375
SEQ. (S) BY (O)	.19311	1.596	3.00	20.00	.222
GROUP (G) BY (O)	.41533	1.178	9.00	66.00	.323
S BY G BY O	.17955	0.467	9.00	66.00	.892
		Item 4			
ORDER (O)	.04378	0.305	3.00	20.00	.821
SEQ. (S) BY (O)	.15138	1.189	3.00	20.00	.339
GROUP (G) BY (O)	.67184	2.116	9.00	66.00	.040*
S BY G BY O	.22248	0.587	9.00	66.00	.803
		Item 5			
ORDER (O)	.37699	4.034	3.00	20.00	.021*
SEQ. (S) BY (O)	.24074	2.114	3.00	20.00	.131
GROUP (G) BY (O)	.70624	2.258	9.00	66.00	.029*
S BY G BY O	.28746	0.777	9.00	66.00	.638
		Item 6			
ORDER (O)	.30919	2.984	3.00	20.00	.056
SEQ. (S) BY (O)	.33946	3.426	3.00	20.00	.037
GROUP (G) BY (O)	.32819	0.901	9.00	66.00	.530
S BY G BY O	.49830	1.461	9.00	66.00	.181

*$p < (.01)df$.

and 5, the Group by Order effects are significant. For these items, the group differences vary by order and the confounded missions.

Table E-7 presents post hoc comparisons between groups for Items 1-3 and 6. Following the Bonferroni method (Kepple, 1982), these tests used an adjusted $\alpha = (.03)/6 = .005$. For Items 1-3, only comparisons with the AOBC NCO group were significant,

Table E-7

Pairwise Comparisons Between Groups for Questionnaire A Items

Group Comparisons	Diff.	Std. Error	t(22)	p
Item 1				
1 vs 2	0.6750	0.2855	2.365	.0272
1 vs 3	0.4875	0.2450	1.919	.0680
1 vs 4	-0.4604	0.2649	-1.738	.0962
2 vs 3	-0.1875	0.2914	-0.644	.5262
2 vs 4	-0.1354	0.3009	-3.773	.0010*
3 vs 4	-0.9479	0.2712	-3.495	.0020*
Item 2				
1 vs 2	0.7750	0.3392	2.285	.0324
1 vs 3	0.1812	0.3018	0.600	.5546
1 vs 4	-0.1646	0.3147	-3.700	.0012*
2 vs 3	-0.5938	0.3462	-1.715	.1004
2 vs 4	-1.9396	0.3575	-5.425	.0000*
3 vs 4	-1.3458	0.3222	-3.556	.0018*
Item 3				
1 vs 2	0.7312	0.2453	2.981	.0068
1 vs 3	0.3562	0.2183	1.632	.1170
1 vs 4	-0.6646	0.2277	-2.919	.0080
2 vs 3	-0.3750	0.2504	-1.498	.1584
2 vs 4	-1.3958	0.2586	-5.398	.0000*
3 vs 4	-1.0208	0.2331	-4.379	.0002*
Item 6				
1 vs 2	0.4812	0.3302	1.457	.1592
1 vs 3	-0.2062	0.2938	-0.702	.0490
1 vs 4	0.5438	0.3064	1.775	.0898
2 vs 3	-0.6875	0.3370	-2.040	.0336
2 vs 4	0.0625	0.3481	0.180	.8588
3 vs 4	0.7500	0.3138	2.390	.0258

Note. Group codes are: AOAC Officers = 1, AOBC Officers = 2, ANCOC NCOs = 3, AOBC NCOs = 4.
*$p < .005$.

confirming that this group has means that are reliably larger for these items. No comparisons were significant for Item 6.

Table E-8 shows the group means at each order for Items 4 and 5. The differences among the groups vary across orders, corresponding to the Group by Order interactions noted before in Table E-6. For Item 4, the means decrease across orders for the AOBC Officers and increase across orders for the AOBC NCOs. The means show no consistent changes across orders for the remaining two groups. These differences are significant for the linear part of the Group by Order interaction, $F(3,22) = 3.77$, $p = .025$. For Item 5, the trend has a U-shaped pattern for the AOAC Officers, but forms an inverted U-shape pattern for the AOBC Officers. No simple patterns appear for the other groups. The significant quadratic part of the Group by Order interaction confirms the form of the different trends among groups, $F(3.22) = 3.59$, $p = .030$. Confounding between orders and mission types prevents drawing conclusions about which variable caused the differences. These interactions did not require more detailed analysis, since unambiguous interpretation of specific group differences or trends is not possible.

Table E-8

Group Means by Order and Missions Type for Items 4-5

	Order and Mission Type			
Instructor Group	1 TRM	2 FOF	3 ATK	4 DEF
	Item 4			
AOAC Officers	3.350	3.225	3.125	3.350
AOBC Officers	2.375	2.625	2.125	1.625
ANCOC NCOs	3.000	2.750	3.000	3.125
AOBC NCOs	3.500	3.750	4.208	3.917
	Item 5			
AOAC Officers	3.875	3.450	3.175	3.775
AOBC Officers	3.250	3.875	3.750	3.125
ANCOC NCOs	3.500	2.500	2.875	2.750
AOBC NCOs	4.167	4.417	4.083	3.875

Appendix F

Analysis of Responses to Questionnaire B

The participants completed Questionnaire B after seeing all four UPAS demonstrations of exercise replays. The participants also had completed Questionnaire A four times, once after each replay. Questionnaire B items are shown in Appendix C. Item 1 on Questionnaire B asked which method(s) they would like to have for use in AARs: the Animated Replay, the Slide Show, both, or neither. As Figure 8 shows, 90% of the participants wanted at least one replay method, and a 63.3% majority wanted both methods. Analysis of the Item 1 response frequencies used the SPSS/PC+ HILOGLINEAR program (Norusis & SPSS, 1990b). The independent variables, Sequences and Groups, formed a three-way contingency table with the third variable, Response, defined by the four response categories. Table F-1 presents the results of the analysis. Tests of partial association showed that the response distributions were not significantly different between sequences or instructor groups. The only significant test involved the marginal distribution for the response categories. Thus, the hypothesis of equal response probabilities among categories was rejected. This result implies that the larger number of "both" responses did not result from chance variation, but reflects a higher probability of response in that category. The goodness-of-fit test statistic for the reduced one-variable log-linear model was $\chi^2(28) = 24.474$, $p = .656$. This result indicates that the residual deviations from the model did not significantly exceed those expected from chance variations.

Table F-1

Analysis of Questionnaire B Item 1 Response Distributions

Effect Name	df	Partial χ^2	p
RESPONSE	3	19.767	.0002*
SEQUENCE	1	0.133	.7149
GROUP	3	0.671	.8801
RESPONSE BY SEQUENCE	3	7.989	.0462
RESPONSE BY GROUP	9	12.923	.1661
SEQUENCE BY GROUP	3	1.516	.6785

*$p < .03$.

Item 5 asked the participants to estimate, on a scale of 0-100% of AARs, how often they would use each replay method. Figure 9 shows how usage estimates for the animated rerun and slide show methods depended on the responses to Item 1. A multivariate analysis of variance (MANOVA) performed on these two

estimates used the SPSS/PC+ procedure. Item 1 Response served as the independent variable with the method estimates treated as repeated-measure variables. Estimated values replaced three missing values in these data. Table F-2 presents the results of the analysis.

The Between-Subjects part of Table F-2 presents the analysis based on the sums of the estimates, while the Within-Subjects part presents the analysis based on the differences between method estimates. The method difference was tested as an *á priori* comparison in this analysis, but was not significant. The statistically significant test for Item 1 Response mainly reflects the high mean usage estimates for both methods given by participants who indicated that they want to use both, compared to low estimates for both methods given by participants who indicated they want to use neither method. The significant Response by Method interaction effect reflects the difference in estimates given by participants who chose only the animated rerun method, or only the slide show method. In both instances, the mean usage estimates were high for the method the participants wanted, compared to low means for the other unwanted method.

Table F-2

Analysis of Questionnaire B Item 5 Estimates of Replay Method Use

Source of Variation	SS	df	MS	F	p
	Between-Subjects (Sum)				
ITEM 1 RESPONSE	12298.00	3	4099.33	6.43	.002*
WITHIN CELLS	15939.93	25	637.60		
	Within-Subjects (Difference)				
REPLAY METHOD	486.75	1	486.75	1.00	.326
RESPONSE BY METHOD	16840.31	3	5613.44	11.58	.000*
WITHIN CELLS	12119.21	25	484.77		

*$p < .03$.

Items 2 and 3 asked the participants to compare AARs using each UPAS replay method to what they usually do in AARs. These items used a 7-point response scale, ranging from "much worse" to "much better." One participant did not answer these two questions, reducing the sample size for these data. The MANOVA shown in Table F-3 used Sequence and Group as independent variables, and treated the two items as repeated-measure variables. Like the previous MANOVA, the analysis is based on

the sum and difference in responses to the two items. The method difference also was tested as an *á priori* comparison in this

Table F-3

Analysis of Questionnaire B Items 2 and 3 Responses Comparing AARs with UPAS Replay Methods to Current AARs

Source of Variation	SS	df	MS	F	p
		Between-Subjects (Sum)			
SEQUENCE (S)	2.03	1	2.03	0.46	.507
GROUP (G)	32.11	3	10.70	2.40	.096
S BY G	2.17	3	0.72	0.16	.920
WITHIN CELLS	93.52	21	4.45		
		Within-Subjects (Difference)			
REPLAY METHOD (M)	.03	1	0.03	0.03	.867
S BY M	1.22	1	1.22	1.04	.319
G BY M	2.27	3	0.76	0.65	.593
S BY G BY M	4.28	3	1.43	1.22	.326
WITHIN CELLS	24.52	21	1.17		

*$p < .03$.

analysis. The statistical tests did not show any effect or interaction to be significant.

For Items 2 and 3 combined, the unweighted mean was 4.369 with a standard error of 0.287. The mean value represents a point between the responses "about equal" and "slightly better." Comparing this mean to the neutral midpoint on the 7-point scale (i.e., a value of 4.0 corresponding to the response "about equal"), the difference was not significant, $t(21) = 1.285$, $p = .213$. The 95% confidence interval for the mean extended from 3.772 to 4.965.

Item 4 and the nearly identical item in Part C directly asked the participants which UPAS replay method they preferred. These items used a 7 point scale ranging from "much prefer" (the animated rerun) to "much prefer (the slide show), corresponding to scale values of 7.0 and 1.0, respectively. One missing response was replaced by an estimated value in these data. The MANOVA was like that for Items 2 and 3, using Sequence and Group as independent variables, treating the two items as repeated-measure variables, and basing the analysis on the sum and difference in responses. In this case, the difference indicates

a change in preference between Items 4 and the Part C item, presumably resulting from seeing intervening items that reminded the participants about specific features of each method, and time factors involved in AARs. The change difference was tested as an

Table F-4

Analysis of Questionnaire B Items 4 and Part C Responses Expressing Preference Between UPAS Replay Methods

Source of Variation	SS	df	MS	F	p
	Between-Subjects (Sum)				
SEQUENCE (S)	4.34	1	4.34	0.62	.440
GROUP (G)	33.78	3	11.26	1.61	.217
S BY G	1.04	3	0.35	0.05	.985
WITHIN CELLS	154.27	22	7.01		
	Within-Subjects (Difference)				
CHANGE (C)	5.66	1	5.66	4.49	.046*
S BY C	.26	1	0.26	0.20	.656
G BY C	3.08	3	1.03	0.81	.499
S BY G BY C	.41	3	0.14	0.11	.954
WITHIN CELLS	27.71	22	1.26		

*$p < .05$.

á priori comparison in this analysis. Table F-4 shows the analysis results. Only the change effect was found to be significant.

For Item 4 the mean response was 3.908, and for the Part C item the mean was 3.273. Thus the participants initially showed little average preference between methods, but later shifted to a small preference for the slide show method. The difference of 0.635 had a standard error of .300. A 95% confidence interval extended from 0.014 to 1.257. Figure 10 has shown the change effect in terms of the percentage of participants that chose a response indicating any degree of preference (a) for the animated rerun method, or (b) for the slide show method.

Items 6 through 13 asked the participants whether they liked or disliked particular aspects of the two replay methods. For both methods, one question concerned each of four instructional characteristics: (a) operating procedures (OP, Items 6 and 7), (b) operating time (OT, Items 8 and 9), (c) type of control (TC), Items 10 and 11), and (d) information flow (IF, Items 12 and 13).

The participants responded using a 7-point scale ranging from "strongly dislike" to "strongly like." These items had no missing responses. The MANOVA for these items, shown in Table F-5, used Sequences and Groups as independent variables. The four characteristics served as multiple dependent variables, with the item pairs related to each feature treated as repeated measures.

Table F-5

Tests of Pillai's Trace Statistic for Questionnaire B Items 6-13

Source of Variation	Trace	F	df	Error df	p
		Between-Subjects (Sums)			
SEQUENCE (S)	.63721	8.343	4.00	19.00	.000*
GROUP (G)	.51609	1.091	12.00	63.00	.383
S BY G	.65022	1.453	12.00	63.00	.167
		Within-Subjects (Differences)			
METHOD (M)	.41442	3.362	4.00	19.00	.031*
S BY M	.02433	0.118	4.00	19.00	.974
G BY M	.67372	1.520	12.00	63.00	.141
S BY G BY M	.37886	0.759	12.00	63.00	.689

*$p < .04$.

This MANOVA design amounted to jointly performing four analyses like the previous ones done on item pairs from Questionnaire B, based on the sums and differences for all four item pairs. In this case, the item differences corresponded to the differences between replay methods, and were tested as *á priori* comparisons.

The multivariate tests showed significant effects for the Sequence and Method sources. Univariate tests of the sequence effect for each dependent variable were significant only for OP, $F(1,22) = 29.219$, $p = .016$. The unweighted mean was 5.156 for subgroups given the UPAS Planned View Display demonstrations with the animated rerun, slide show, slide show, animated rerun method sequence, ccmpared to 3.177 for the subgroups given the slide show, animated rerun, animated rerun, slide show sequence. The difference of 1.979 had a standard error of .366. Participants who saw the former sequence apparently developed a greater liking for the UPAS Planned View Display replay procedures regardless of method, whereas those who saw the latter sequence came to dislike the procedures. The reason for this difference is obscure, since related sequence effects did not appear in prior analyses of Questionnaire B items, or for other items in this analysis.

Sequence effects regarding responses to Questionnaire B items would be expected to produce general effects, and not be related to only one aspect of the UPAS Planned View Display replays.

The main comparisons of interest in this analysis were the differences between the animated rerun and slide show exercise replay methods. Table F-6 shows the method means for each

Table F-6

Unweighted Means for UPAS Exercise Replay Methods

	Instructional Characteristic			
Method	OP	OT	TC	IF
Animated Rerun	3.490	3.888	5.862	5.069
Slide show	4.844	4.808	5.879	5.285

Note. OP = Operating Procedures, OT = Operating Time, TC = Type of Control, IF = Information Flow.

instructional characteristic. Both OP and OT means have sizable differences between methods. In both cases, the differences show that the participants liked the characteristics of the slide show method better than the animated rerun method. On the other hand, the TC and IF means show small differences between methods. Tests of significance and confidence intervals for the differences are presented in Table F-7. The differences were significant only for the OP and OT characteristics.

For the TC and IF characteristics, means for both methods are greater than the scale neutral point (4.0), indicating that the participants rated both methods favorably. The combined mean for TC was 5.871, significantly larger than the neutral point, $t(22) = 9.154$, $p = .000$. The standard error for this mean was .196, and a 95% confidence interval ranged from 5.463 to 6.279. The combined mean for IF was 5.177, significantly larger than the neutral point, $t(22) = 6.089$, $p = .000$. The standard error was .193, and the 95% confidence interval was from 4.776 to 5.578.

In order to determine the best-liked characteristics for each replay method, additional analyses examined differences among item means. Table F-8 presents *post hoc* pairwise item comparisons between means. For the animated rerun method, the means for Items 11 and 13 (related to the IF and animated rerun characteristics) were significantly larger that the means for Items 7 and 8 (related to OP and OT). For the slide show method,

Table F-7

Significance Tests and Confidence Intervals on Contrasts Between Animated Rerun and Slide Show Methods of Presentation for Items on Questionnaire B Paired by Instructional Characteristic

Item Pair	Diff.	SE	t(22)	p	Lower-95% CL-Upper	
7-6, OP	-1.354	0.412	-3.287	.003*	-2.209	-0.500
8-9, OT	-0.921	0.404	-2.276	.033*	-1.760	-0.817
11-10, TC	-0.017	0.287	-0.058	.954	-0.613	0.580
13-12, IF	-0.217	0.530	-0.409	.687	-1.316	0.883

Note. OP = Operating Procedures, OT = Operating Time, TC = Type of Control, IF = Information Flow.
*p < .05.

Table F-8

Pairwise Comparisons Among Questionnaire A Items 6-13

Item Comparisons	Diff.	SE	t(22)	p
Animated Rerun				
7 vs 8	-0.3979	0.215	-1.847	.078
7 vs 11	-2.3792	0.284	-8.365	.000*
7 vs 13	-1.5792	0.350	-4.514	.000*
8 vs 11	-1.9750	0.288	-6.847	.000*
8 vs 13	-1.1812	0.360	-3.279	.003*
11 vs 13	0.7937	0.392	2.023	.055
Slide Show				
6 vs 9	-0.0	0.378	-0.926	.926
6 vs 10	1.0	0.333	3.113	.005*
6 vs 12	0.4417	0.384	1.150	.263
9 vs 10	1.0708	0.296	3.623	.002*
9 vs 12	0.4771	0.407	1.172	.254
10 vs 12	-0.5937	0.243	-2.449	.023

*p < .005

the Item 10 mean (related to TC) was significantly greater than the means for Items 6 and 9 (related to OP and OT). None of these three means differed significantly from the Item 12 mean (related to IF).

Appendix G

Analysis of Responses to Questionnaire C

After completing Questionnaire B, the participants were shown examples of UPAS displays intended to supplement the Planned View Display replays as AAR aids. These UPAS displays included the (a) Exercise Timeline, (b) Battle Flow, (c) Battle Snapshot, (d) Snapshot with Line-of-Sight, and (e) Graphs. One or more examples of each type of display were prepared for the same four mission exercises used to demonstrate the Planned View Display replays. The participants completed Questionnaire C after seeing slide shows that presented all the display examples.

The first item asked which displays the participants wanted to have in UPAS for use in AARs. The responses were "yes", "neutral", or "no". One participant who gave no answer for three displays was omitted from the analyses performed on these data. Single missing responses for two participants were counted as "neutral". Table G-1 shows the percentage of "yes" responses for each display separately for officer and NCO instructor groups. The percentages for the two groups appear to differ for the Exercise Timeline, Battle Flow, and line-of-sight displays, but are similar for the Battle Snapshot and GR displays.

Table G-1

Percentage of Armor School Instructors Wanting Types of UPAS Displays for Use in AARs

	Type of Display				
Instructor Group	TL	BF	BS	LS	GR
AOAC & AOBC Officers	13.3	26.6	86.7	46.7	66.7
ANCOC & AOBC NCOs	78.6	57.1	71.4	71.4	64.3

Note. TL = Exercise Timeline, BF = Battle Flow, BS = Battle Snapshot, LS = Snapshot with Line-of-Sight, GR = Graphs.

Analysis of the response frequencies used the SPSS/PC+ HILOGLINEAR procedure (Norusis & SPSS, 1990b). The analysis was performed on two-way tables formed for each display by combining the Group independent variable with the Response dependent variable. Results of the maximum likelihood tests of independence are shown in Table G-2. These tests used $\alpha = .03$ based on the three degrees of freedom for the independent variable. Significant differences among the group's response distributions occurred only for the Exercise Timeline display.

Despite appearances, the differences for Battle Flow and line-of-sight were not significant.

Table G-2

Loglinear Tests of Independence Between Groups and Responses for UPAS Display Types

Display Type	df	Partial χ^2	p
Exercise Timeline	6	17.328	.0082*
Battle Flow	6	10.903	.0914
Battle Snapshot	6	7.080	.3135
Snapshot with Line-of Sight	6	9.804	.1332
Graphs	6	3.747	.7108

*$p < .03$.

Group contrasts were then examined using the LOGLINEAR procedure. The goodness of fit for the two-way table was acceptable if the officer-NCO contrast was kept in the loglinear model, but the contrasts between the officer groups and between the NCO groups were removed, $\chi^2(4, N = 29) = 3.757$, $p = .440$. Therefore, the response distributions only differed between the officer and NCO groups.

Display differences were tested separately for the officer and NCO groups using the SPSS/PC+ procedure NPAR TESTS. The Friedman rank-order test for related samples was done with $\alpha = .04$, using rank values to replace the three ordinal response categories. The test was significant for the officers, $\chi^2(4, N = 15) = 19.893$, $p = .0005$, but not for the NCOs, $\chi^2(4, N = 14) = 0.457$, $p = .9775$. These results indicate that the officer group differed about wanting the displays available in UPAS, while the NCO group tended to approve all the displays about equally. For the officer group, the Wilcoxon matched-pairs test was used to compare the five displays pairwise (i.e., ten comparisons) with $\alpha = (.04)/10 = .004$. These tests showed that the Exercise Timeline was wanted less than the Battle Snapshot, $z = 3.059$, $p = .0022$. All the other display differences were statistically insignificant.

A second part of the first item in Questionnaire C asked the participants to rank order priorities among the displays according to their preference for keeping them in the UPAS. This measure produced very similar results to the yes-no question. The mean ranks are given in Table G-3, with lower rank values representing higher priorities. One officer and one NCO did not provide rankings, and do not contribute to these means. For the Table G-3

Mean Priority Ranking by Armor School Instructors for Keeping Types of Displays in UPAS to Use in AARs

	Type of Display				
Instructor Group	TL	BF	BS	LS	GR
AOAC & AOBC Officers	4.29	3.71	1.79	2.64	2.57
ANCOC & AOBC NCOs	2.50	3.14	2.64	3.43	3.29

Note. TL = Exercise Timeline, BF = Battle Flow, BS = Battle Snapshot, LS = Snapshot with Line-of-Sight, GR = Graphs.

officers, the Battle Snapshot was given the highest priority, while the Exercise Timeline was given lowest priority. For the NCOs, the differences in priorities among displays were relatively small. The mean ranks for the Exercise Timeline display show the largest difference between groups.

Using the NPAR TESTS procedures, rank tests were done on the differences among instructor groups for each display type. None of the Kruskal-Wallis tests were statistically significant, as shown in Table G-4. However, the comparison between officer and NCO groups for the Exercise Timeline display was significant with $\alpha = .01$. The Mann-Whitney test statistic was $U = 42.5$, $z = -2.682$, $p = .007$ (corrected for ties). Similar group comparisons were not significant for the other displays.

Table G-4

Kruskal-Wallis Rank Tests of Differences in Priorities Among Groups for UPAS Display Types

Display Type	*df*	χ^2	*p*
Exercise Timeline	4	8.266	.0408
Battle Flow	4	3.714	.2941
Battle Snapshot	4	2.828	.4189
Line-of-Sight	4	3.925	.2697
Graphs	4	1.538	.6736

Note. *N* = 28 for each test.

Tests of display differences were done within the officer and NCO groups using NPAR TESTS procedures (Norusis & SPSS,

1990a). The Friedman rank-order test was significant (α = .04) for the officers, $\chi^2(4, N = 14) = 22.114$, $p = .0002$, but not for the NCOs, $\chi^2(4, N = 14) = 3.714$, $p = .4461$. For the officers, pairwise Wilcoxon tests (α = .004) showed that the Exercise Timeline had lower priority (larger rank values) than both the Battle Snapshot display, $z = -2.900$, $p = .0037$, and the line-of-sight display, $z = -3.045$, $p = .0023$. Also, the Battle Flow display had lower priority than the Battle Snapshot display, $z = -3.180$, $p = .0015$. Other display comparisons were significant.

Items 2-6 on Questionnaire C asked the participants to rate the usefulness of the five supplementary UPAS displays for an AAR in relation to different types of missions. The four missions were (a) Tactical Road March (TRM), (b) Move/React to Contact (MTC), (c) Attack (ATK), and (d) Defense (DEF). The ratings used a 5-point scale with ordered categories "not...", "slightly...", "somewhat...", "fairly...", and "very useful". One nonresponse required the use of Yates' method to estimate the missing value.

An initial multivariate analysis with SPSS/PC+ procedure MANOVA (Norusis & SPSS, 1990b) included Sequence and Group as between-subject variables and Display and Mission as repeated-measure variables. No effects or interactions involving the Sequence variable were significant in this analysis, so subsequent analyses dropped this variable. The Group main effect was not significant in the between-subjects part of the simplified analysis, $F(3,26) = 2.23$, $p = .109$. The multivariate test results are given in Table G-5. The significant interaction effects show that the rated usefulness of the UPAS displays varies by Group, and also by Mission.

Table G-5

Tests of Pillai's Trace Statistic for Repeated-Measure Effects

Source of Variation	Trace	F	df	Error df	p
Display (D)	.45716	4.842	4	23	.006*
Mission (M)	.38033	4.910	3	24	.008*
D BY M	.69764	2.884	12	15	.028*
GROUP (G) BY D	.77374	2.172	12	75	.022*
GROUP (G) BY M	.19778	0.612	9	78	.784
G BY D BY M	1.27100	1.041	36	51	.441

*$p < (.01)df$.

Table G-6 shows the ratings of display usefulness for each instructor group. The ratings exhibit officer-NCO differences for displays similar to those found with the yes-no questions, and for the priority rankings. Analyses of the differences

between groups for each display type showed significant effects only for the Exercise Timeline, $F(3,26) = 5.01$, $p = .007$, and for the Battle Flow, $F(3,26) = 4.43$, $p = .012$. Tests of contrasts between groups for these displays showed that the differences between officers and NCOs, as presented in Figure 22, accounted for the major part of the group variation. For the Exercise Timeline display, $t(26) = 3.869$, $p = .001$, and for the Battle Flow display, $t(26) = 3.589$, $p = .001$.

Table G-6

Mean Usefulness Ratings for Types of Displays by Instructor Group

	Type of Display				
Instructor Group	TL	BF	BS	LS	GR
AOAC Officers	2.250	2.222	4.444	3.972	3.444
AOBC Officers	1.792	2.042	3.417	2.625	3.042
ANCOC NCOs	3.562	3.750	3.250	3.875	3.218
AOBC NCOs	3.714	3.500	4.000	3.786	3.786

Note. TL = Exercise Timeline, BF = Battle Flow, BS = Battle Snapshot, LS = Snapshot with Line-of-Sight, GR = Graphs.

Table G-7 shows the ratings of display usefulness for each type of mission. In order to isolate the differences that produced the Display by Mission interaction, MANOVA analyses were done separately on the data for each display. As shown in Table G-8, the mission differences were significant with the Battle Snapshot, line-of-sight, and GR displays, and was nearly significant with the Exercise Timeline display. For these Exercise Timeline, line-of-sight, and GR displays, contrasts comparing the average for the platoon-level missions (TRM and MTC) with that for the company-level missions (ATK and DEF) were significant ($\alpha = .01$). These contrast tests were: (a) for the Exercise Timeline display, $t(26) = 2.894$, $p = .008$, (b) for the line-of-sight display, $t(26) = 4.009$, $p = .000$, and (c) for the GR display , $t(26) = 4.124$, $p = .000$.

Pairwise comparisons given in Table G-9 indicate that the significant platoon-company level contrasts for the line-of-sight and GR displays result from smaller means for the TRM mission alone. A similar pattern of pairwise differences was found for the Battle Snapshot display, although the TRM-DEF comparison failed to be significant. The Exercise Timeline display means do not follow this pattern, and the contrast appears to reflect a real difference between levels.
Table G-7

Mean Usefulness Ratings for Types of Displays by Mission

	Type of Display				
Mission	TL	BF	BS	LS	GR
Tactical Road March	3.160	2.846	3.235	2.729	2.589
Move/React to Contact	2.978	2.965	4.034	3.742	3.462
Attack	2.660	2.961	4.030	3.778	3.669
Defense	2.520	2.742	3.812	4.003	3.741

Note. TL = Exercise Timeline, BF = Battle Flow, BS = Battle Snapshot, LS = Snapshot with Line-of-Sight, GR = Graphs.

Table G-8

Tests of Pillai's Trace Statistic for the Mission Effect with Each Type of Display

Type of Display	Trace	F	df	Error df	p
Timeline	.29920	3.415	3	24	.034
Battle Flow	.16303	1.558	3	24	.225
Battle Snapshot	.34948	4.298	3	24	.015*
Line-of-Sight	.47021	7.100	3	24	.001*
Graphs	.45801	6.760	3	24	.002*

*$p < .03$.

Table G-9

Pairwise Comparisons Between Mission Means for Types of Displays

Group Comparisons	Diff.	Std. Error	t(22)	p
Timeline				
TRM vs MTC	0.2000	0.169	1.18	.246
TRM vs ATK	0.5333	0.213	2.50	.018
TRM vs DEF	0.6667	0.260	2.57	.016
MTC vs ATK	0.3333	0.111	3.01	.005*
MTC vs DEF	0.4667	0.178	2.63	.014
ATK vs DEF	0.1333	0.104	1.28	.211

(Continued)

Table G-9 (Continued)

Pairwise Comparisons Between Mission Means for Types of Displays

Group Comparisons	Diff.	Std. Error	$t(22)$	p
Battle Flow				
TRM vs MTC	-0.1000	0.246	-0.41	.687
TRM vs ATK	-0.1000	0.268	-0.37	.712
TRM vs DEF	0.1000	0.237	0.42	.676
MTC vs ATK	0.0000	0.107	0.00	1.000
MTC vs DEF	0.2000	0.130	1.53	.136
ATK vs DEF	0.2000	0.101	1.99	.056
Battle Snapshot				
TRM vs MTC	-0.8000	0.206	-3.89	.001*
TRM vs ATK	-0.8000	0.222	-3.61	.001*
TRM vs DEF	-0.6000	0.212	-2.83	.008
MTC vs ATK	0.0000	0.107	0.00	1.000
MTC vs DEF	0.2000	0.169	1.18	.246
ATK vs DEF	0.2000	0.101	1.99	.056
Line-of-Sight				
TRM vs MTC	-1.0667	0.239	-4.46	.000*
TRM vs ATK	-1.1000	0.241	-4.56	.000*
TRM vs DEF	-1.3000	0.263	-4.94	.000*
MTC vs ATK	-0.0333	0.131	-0.25	.801
MTC vs DEF	-0.2333	0.164	-1.42	.165
ATK vs DEF	-0.2000	0.101	-1.99	.056
Graphs				
TRM vs MTC	-0.9000	0.241	-3.73	.001*
TRM vs ATK	-1.1667	0.272	-4.30	.000*
TRM vs DEF	-1.2000	0.269	-4.47	.000*
MTC vs ATK	-0.2667	0.135	-1.97	.058
MTC vs DEF	-0.3000	0.137	-2.19	.037
ATK vs DEF	-0.0333	0.033	-1.00	.326

*$p < .005$.

Appendix H

"Think-Aloud" Protocols Regarding the UPAS Materials

The purpose of this interview is to help make the UPAS materials better suited to fit your needs. We are not interested in evaluating your instructional methods, but rather in better understanding your views on the visual aids needed to enhance your AARS. As you go through this exercise, please share with me the following issues as soon as they become evident to you:

1. the instructional point or points that must be made.
2. the instructional purpose(s) associated with that point(s).
3. the visual-aid you need to help you make your point with the student.

The visual-aid can be any or all of the following UPAS displays (show them the display). It could also be a visual display which is not available in the current UPAS system. Occasionally, I will prompt you to either acquire this information or help me better understand your ideas. I may also ask you some follow-up questions after this exercise is complete and before the AAR has started. Again, the purpose of such questions is to better help us understand your ideas. Thank you for your cooperation in this matter.

INSTRUCTOR_____ CLASS_____ DATE_____

EXERCISE TYPE_____

START TIME_____ END TIME_____

EVENT_____ TIME_____

INSTRUCTIONAL POINT:

INSTRUCTIONAL PURPOSE:

PVD DISPLAY_____ BATTLEFLOW _____ SNAPSHOT/LS_____
TIMELINE: R B_____ GRAPh_____ TYPE_____

ENHANCEMENTS OR OTHER (DESC):

REASONS FOR CHOICE(S):