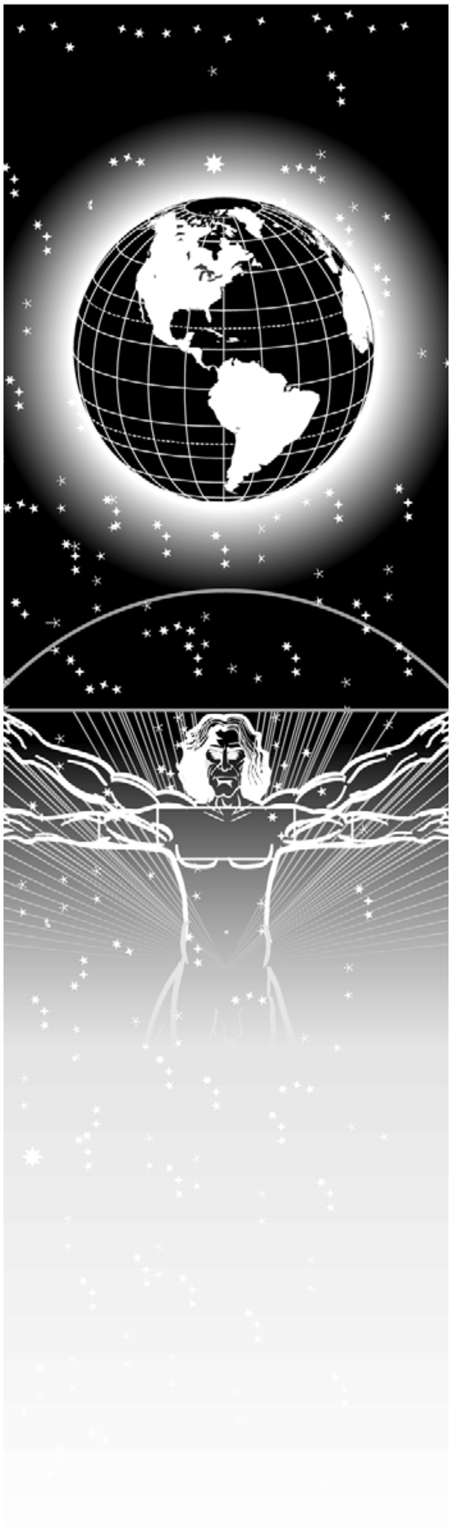
AFRL-HE-AZ-TP-2003- 0002



UNITED STATES AIR FORCE RESEARCH LABORATORY

PRELIMINARY EVALUATION OF THE CONTROL INTERFACE FOR A TRAINING SIMULATION SYSTEM

Joseph Lloyd Weeks
Air Force Research Laboratory
Warfighter Training Research Division
6030 South Kent Street
Mesa, AZ 85212-6061

L. Bruce McDonald
McDonald Research Associates
120 University Park Drive, Suite 200
Winter Park, FL 32792

September 2003

AIR FORCE MATERIEL COMMAND
AIR FORCE RESEARCH LABORATORY
Human Effectiveness Directorate
Warfighter Training Research Division
6030 South Kent Street
Mesa AZ 85212-6061

NOTICES

JOSEPH L. WEEKS
Project Scientist

HERBERT H. BELL
Technical Advisor

CURTIS J. PAPKE, Colonel, USAF
Chief, Warfighter Training Research Division

Direct requests for copies of this report to:

Defense Technical Information Center
8725 John J. Kingman Road, Suite 0944
Ft. Belvoir, VA 22060-6218
http:// stinet.dtic.mil

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
September 2003

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
May 2002 – May 2003

4. TITLE AND SUBTITLE
Preliminary Evaluation of the Control Interface for a Training Simulation System

5a. CONTRACT NUMBER
F41624-97-D-5000

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
63227F

6. AUTHOR(S)
Joseph L. Weeks
L. Bruce McDonald

5d. PROJECT NUMBER
4924

5e. TASK NUMBER
B0

5f. WORK UNIT NUMBER
02

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

McDonald Research Associates
120 University Park Drive, Suite 200
Winter Park, FL 32792

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division
6030 South Kent Street
Mesa AZ 85212-6061

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL; AFRL/HEA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-HE-AZ-TP-2003-0002

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

13. SUPPLEMENTARY NOTES This technical paper documents a presentation at the North Atlantic Treaty Organization, Research and Technology Agency, Systems Concepts and Integration symposium, Critical Design Issues for the Human-Machine Interface, 19-21 May 2003. Air Force Research Laboratory Technical Monitor: Dr. Joseph Lloyd Weeks, Comm: (480) 988-6561, Ext. 249, DSN: 474-6249

14. ABSTRACT
The United States Air Force Research Laboratory is conducting research and development of a computer-based, simulation capability to support training in decision making and team coordination for security forces ground operations. The simulation capability allows an instructor to start and control a simulation exercise on trainee computer workstations connected via a local area network. Simulation software supports the interaction of trainees with each other and with computer-generated forces (CGFs) that imitate behavior of enemy, neutral, and friendly troops and civilians. For the security forces simulation capability, the goal is to design a control interface that instructors can learn to use in two hours and trainees can learn to use in thirty minutes. To achieve this goal, a Windows-based control interface was adopted as the initial point of departure. Menu options were developed to correspond to the standard mission planning procedures used by security forces and drag-and-drop functions were developed to replace menu options to contribute to usability. This paper describes the emerging control interface, the approach to and outcomes from a field evaluation of the interface to include actual times required for instructors and trainees to learn to use the system and instructors' acceptability evaluations. Summary conclusions shed light on critical human-machine interface design issues for computer-based training simulations.

15. SUBJECT TERMS
Computer-based simulation; Computer-based training; Computer-generated forces; Control interfaces; Decision making; Human-machine interface; Team coordination; Teams; Training

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(include area code)*
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED	20	Ms Liz Casey / 480.988.6561 x-188 DSN 474-6188

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

SYSTEMS CONCEPTS AND INTEGRATION PANEL SYMPOSIUM

Critical Design Issues for the Human-Machine Interface

PRELIMINARY EVALUATION OF THE CONTROL INTERFACE FOR A TRAINING SIMULATION SYSTEM

Dr. Joseph Lloyd Weeks
Warfighter Training Research Division,
United States Air Force Research Laboratory,
6030 South Kent Street, Mesa, AZ, USA 85212-6060
Tel: +1 480 988 6561 Ext 249 Fax: +1 480 988 6285
E-mail: joseph.weeks@williams.af.mil

ABSTRACT

The United States Air Force Research Laboratory is conducting research and development of a computer-based, simulation capability to support training in decisionmaking and team coordination for security forces ground operations. The simulation capability allows an instructor to start and control a simulation exercise on trainee computer workstations connected via a local area network. Simulation software supports the interaction of trainees with each other and with computer-generated forces (CGFs) that imitate behavior of enemy, neutral, and friendly troops and civilians. Radio functions allow multi-channel communication among instructors, trainees, and CGFs. The simulation capability can provide multiple terrain environments without being limited to unchanging terrain features associated with a fixed field-training site. It can support a wide range of alternatives for a given situation so trainees can practice decisionmaking and team coordination in a number of scenarios with varying threat and environmental conditions. The simulation capability is intended to support training in leadership and command and control; it does not support training related to trigger-pulling. The greatest obstacles to incorporating such simulations into formal training are associated support requirements. Typically, on-site technicians are required to design and develop simulation exercises in support of learning objectives and must be present during the instructional event to control the exercise, serve as role players, task CGFs, and provide simulation replays to support after-action reviews. Costs of support requirements are a barrier to fielding simulation technology for formal training; hence, design and development of a simulation control interface that can be directly used by instructors and trainees is an important R&D objective. For the security forces simulation capability the goal is to design a control interface that instructors can learn to use in two hours and trainees can learn to use in thirty minutes. To achieve this goal, a Windows-based control interface was adopted as the initial point of departure. Menu options were developed to correspond to the standard mission planning procedures used by security forces and drag-and-drop functions were developed to replace menu options to contribute to usability. The paper will describe the emerging control interface, the approach to and outcomes from a field evaluation of the interface to include actual times required for instructors and trainees to learn to use the system and instructors' acceptability evaluations. Summary conclusions will shed light on critical human-machine interface design issues for computer-based training simulations.

PRELIMINARY EVALUATION OF THE CONTROL INTERFACE FOR A TRAINING SIMULATION SYSTEM

Dr. Joseph L. Weeks
Warfighter Training Research Division,
United States Air Force Research Laboratory,
6030 South Kent Street, Mesa, AZ, USA 85212-6061
Tel: +1 480 988 6561 Ext 249 Fax: +1 480 988 6285
E-mail: joseph.weeks@williams.af.mil

Dr. L. Bruce McDonald
McDonald Research Associates
120 University Park Drive, Suite 200
Winter Park, FL 32792
Tel: +1 407 644 1900 Fax: +1 407 679 9558
E-mail: bruce@mrassociates.com

INTRODUCTION

United States Air Force security forces are responsible for installation security, airbase defense, and military police services. Whether security forces discover an improvised explosive device in a car or confront representatives of a non-governmental organization at an entry control point, such situations are reported to a command post whether it is called central security control or a law enforcement desk. Once situations are reported, quick and accurate decisions by security forces leaders are critical to handling the situation properly and fundamental to the protection of personnel and assets.

Command post exercises are routinely conducted to train personnel to respond to diverse situations. Nevertheless the command and control of security forces continues to be recognized as a high-emphasis training area (Weeks, Garza, Archuleta, & McDonald, 2001). As a result of established training needs and technology opportunities, the Air Force Research Laboratory is conducting research and development of an interactive simulation capability (Weeks & McDonald, 2002). With the help of security forces, the capability is being tailored to meet training needs. Computer-driven, simulations are regarded as an affordable way to accelerate acquisition of expertise in leadership, decisionmaking, and team coordination; however, transition and fielding of such a capability depends on its usability and proven training effectiveness.

An illustration of the training device is presented at Figure 1. It will consist of standard computers connected by a local area network. To reduce acquisition and maintenance costs standard visual displays will be used. A trainee station will consist of a computer processor, computer monitor, keyboard, mouse, and radio headset and microphone for communications with instructor, other trainees, and computer-generated forces (CGFs). The instructor station will include an additional computer for logging simulation exercises for after-action reviews. In Figure 1, trainees are illustrated as a shift leader and subordinate flight leaders but could alternatively be a flight leader and subordinate squad leaders or a defense force commander, operations officer, and area supervisors. The capability is being developed as a multi-echelon, command and control, training device to support different training requirements.

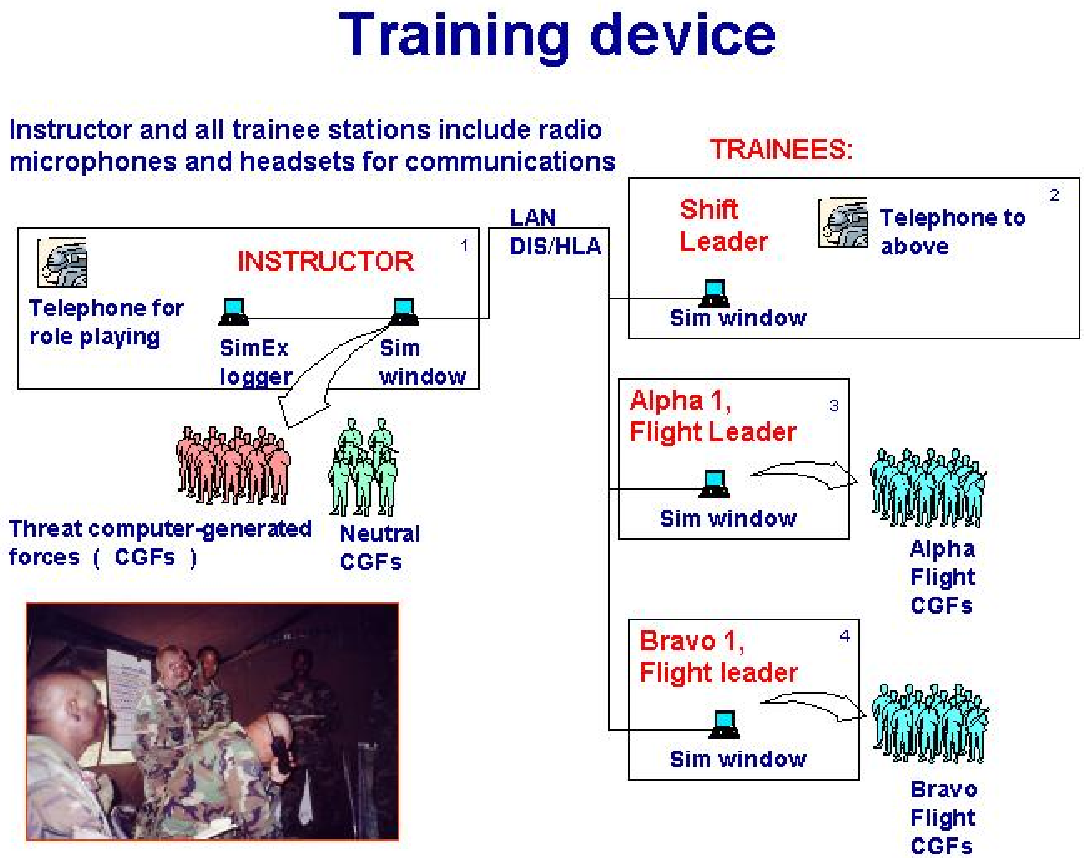


Figure 1. The training device

The capability will support training for security planning and execution of the defense. During application, the instructor would present a digital map identifying the area of operations on the local area network. All trainees would simultaneously view the map, collaborate in conducting a security vulnerability analysis, and annotate the map to develop a security plan by creating and positioning computer models in the area of operations. Computer models are being developed to represent sensors, obstacles, fighting positions, vehicles, communications, and semi-automated, computer-generated forces. CGFs are designed to move, sense, shoot, and communicate. After the security plan is completed, trainees would share the annotated map with the instructor. The instructor would evaluate the plan and could create, position, and task neutral or threat CGFs to present security situations to test the plan. After trainees have developed the plan and the instructor creates and tasks CGFs to present security situations, the exercise could begin. A situation report from a security forces CGF via the radio headset or tasking by the instructor acting as a Wing Operations Center Commander would be the initiating action for trainee decisionmaking and teamwork. Communications from CGFs to trainees and among trainees would be central to evaluation of command and control performance. All communications via radio microphones and actions occurring on the visual display would be logged for after-action reviews.

The simulation capability is currently being refined through research and development. Research objectives include development of a usable control interface, realistic behaviors for CGFs, development of simulation exercises to support learning objectives, and evaluation of system usability, model validity, and training effectiveness. The R&D project includes multiple field evaluations with the participation of end users. The development strategy is to collect evaluation feedback and apply it to refinement of the capability in an effort to accelerate transition to the field. Evaluations will include assessments of the usability of the simulation control interface, assessments of the validity of computer models for sensors, obstacles, weapons, communications, and CGFs, assessments of the acceptability of draft training scenarios, and evaluation of the training effectiveness of the complete training system. The purpose of this paper is to describe the approach, results, and conclusions from a preliminary usability evaluation and to make general recommendations for structuring usability evaluations of the human-machine interface.

APPROACH

Brewer, Armstrong, and Steinberg (2002) state "usability testing … verifies that a system or product design meets our expectations for usefulness and satisfaction before we move into production (p 403)." They define usability as "the degree to which the design of a device or system may be used effectively and efficiently by a human (p 403)" and point out that the important issue in arriving at a definition is how to measure usability so that measurements can be used to improve the design. Brewer, et al. (2002) outline three general approaches to usability evaluation including surveys (using self-report data collection methods), usability inspections (specialists scrutinize a design according to a systematic approach and judge its acceptability against certain criteria), and experimental tests (based on quantifying operator performance using controlled data collection techniques). Usability inspection best describes the approach. The specialists were instructors assigned to the Security Forces 96th Ground Combat Training Squadron (96th GCTS). They were first trained to use the simulation control interface. After familiarization and training, they evaluated the interface on critical criteria, identified problems, and recommended improvements. These specialists were not experts in software or human factors engineering, but they did represent a final authority for usability, the end user.

Because trainees are also end users, it was hoped they would participate by providing baseline training times. However, the field evaluation was scheduled between training classes; therefore trainees were not available. Although trainees are important end users and should participate in the evaluation, instructors indicated that a full course schedule would not have allowed trainees' participation if they had been present. The exclusion of trainees did provide the advantage of giving instructors time to concentrate on the technology evaluation without simultaneously being engaged in training. However, trainee participation would have to be obtained elsewhere.

The complete training system consists of several computers linked in a local area network. However, only one laptop computer was used for the preliminary evaluation. The strategy was to modify the control interface on the basis of change recommendations before taking the complete training system to the field for evaluation.

To describe the usability inspection approach, Brewer, et al (2002) present an example of a computer interface. It is a single dialogue box. The usability issue is whether to position control buttons on the bottom left or bottom right of the dialogue box. Compared to the dialogue box described by Brewer, et al (2002), the interface evaluated here is immense. It consists of over 450 different controls including menus, tools, dialogue boxes, intermediate control windows, and simulation window symbols. One recommended approach to usability testing is based on presenting a control interface to end users, not informing them how to use it, observing if they can deduce how to use it, and requesting feedback concerning improvements (Andre, personal communication, 2003). Although this would be a useful approach for the type of interface that Brewer, et al. (2002) describe, it was not used here. In addition to the great number of interface controls, most instructors had no experience with simulation capabilities like the one evaluated. For them it was novel; so it was not practical to adopt a discovery learning approach. To minimize the duration of the evaluation period and maximize meaningful usability evaluations, it was necessary to familiarize participants with the interface in advance.

The simulation control interface used for the preliminary evaluation was delivered with Version 2.0 of the simulation software and is presented at Figure 2. It consisted of a simulation window in which a digital map is presented; a menu bar consisting of menus, menu options, and menu option labels; a tool bar consisting of tool bar buttons and information windows; a pop-up menu presented in the simulation window and accessed through a right mouse click, a mouse, and a standard computer keyboard. The radio microphone and headset and simulation logger were not included in the preliminary evaluation. These interface devices will be tested during evaluation of the complete training system.

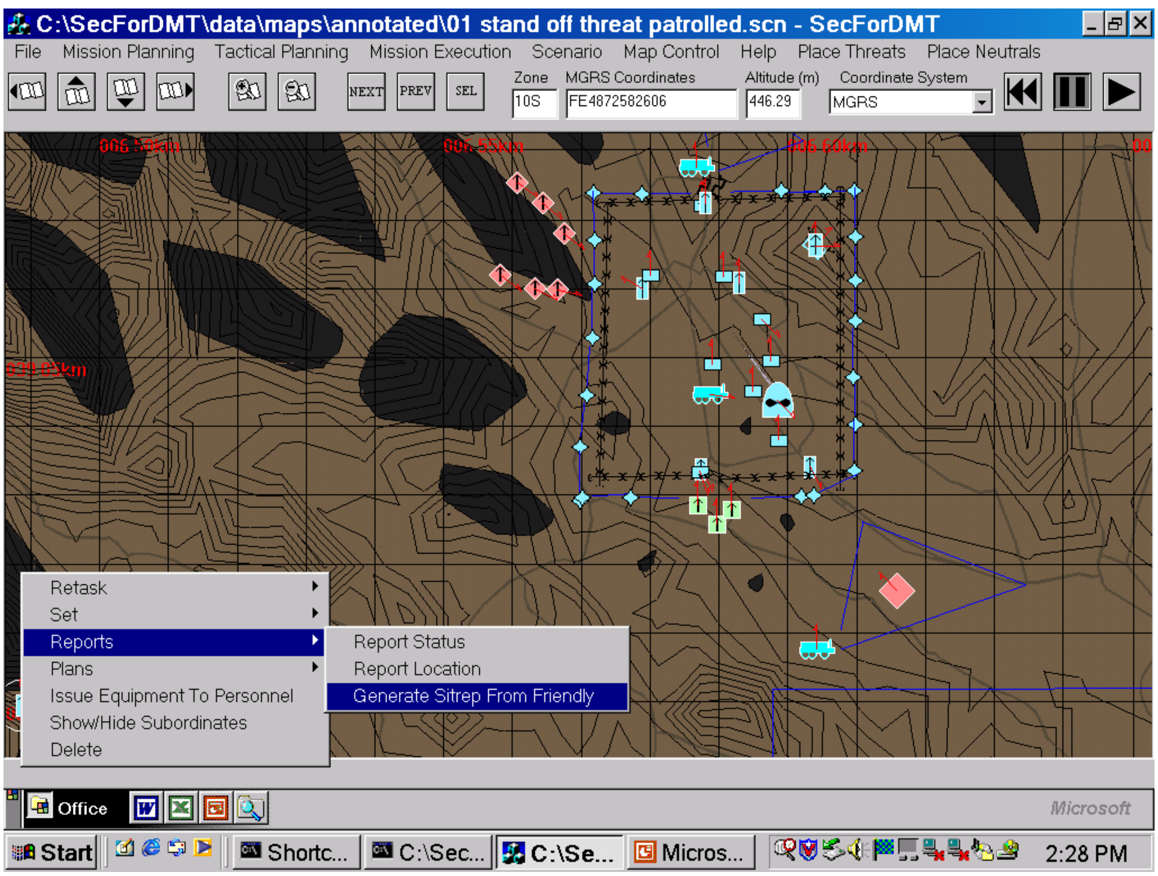


Figure 2. The simulation control interface

One of the greatest obstacles to incorporating such simulations into formal training is the support costs. For training simulation systems currently available, on-site technicians are required to design and develop simulation exercises in support of learning objectives and must be present during the instructional event to control the exercise, serve as role players, task CGFs, and provide simulation replays to support after-action reviews. Costs of support requirements are a barrier to fielding simulation technology for formal training. Thus, development of a control interface that can be directly used by instructors and trainees and that eliminates the need for an on-site technician is an important R&D objective. At the beginning of the project, a usability standard was established and defined as the time required to learn to use the device. The expectation is the shorter the training time; the greater the usability. The intent is to compare the standard with observed and threshold training times as a guide for system modifications. Training time standards for usability are that instructors will be trained to use the device in 2 hours and students will be trained in 30 minutes. Evaluation of training time was one of the objectives of the preliminary evaluation.

The evaluation was conducted one instructor at a time. This approach minimized adverse impact on day-to-day activities of the training group that could have occurred if several instructors were tasked to support the evaluation in mass. Each instructor was trained to use the control interface illustrated in Figure 2 and their training time was recorded. It was explained to each instructor that they were not being evaluated; rather their training time was being recorded to derive an estimate of observed training time for evaluating the device. Threshold training time was defined as the maximum time that could be allocated for an instructor to be trained to use the device. After training on all tasks, instructors were asked to estimate threshold training time.

The simulation control interface included over 450 control options. Rather than attempting to train instructors on all control options, a sample of tasks was selected to represent those instructors would perform during a simulation exercise. Task selection involved a trade-off between practicality and an evaluation of all control options. A balance was sought between what participants might regard as an intolerably long evaluation period and evaluation of all control options. The original task list consisted of 24 tasks in three categories representing control of the simulation window, placing resources, and tasking CGFs. The 24 tasks and associated control options are a non-random sample from the population of tasks. Rather than randomly sampling and presenting tasks, they were carefully selected and sequenced for meaningfulness. If more than 24 tasks had been selected for the evaluation, total observed training time would have been greater. However, the more tasks selected; the greater the length of the evaluation period, and the less likely it would be to obtain instructors' willing cooperation and meaningful evaluation input. The 24 tasks and control options used for the preliminary evaluation are described in the Appendix.

To begin the evaluation, each instructor was presented a briefing describing the complete training system and training concept. After the briefing, they were asked to read and sign a disclosure and consent form and to complete a background questionnaire to obtain information about their time in service, training, and experience with computers. Measures of training time were obtained separately for each instructor. They were told how to perform each task, showed how to do it, and asked to independently perform the task with assistance. They were asked to indicate when they had learned to use interface controls for the task. When each instructor stated she or he had learned to perform the task, training time was declared complete. Immediately after training for each task, the instructor was asked to perform the task independently. It was noted whether it was performed with or without assistance. It was assumed that if instructors satisfactorily performed the task without assistance, they had learned to use the controls. If the instructor asked for assistance, recorded training time for that task for that instructor was doubled. Task training time was cumulated over all tasks to obtain an estimate of training time for a single instructor and averaged over all instructors to estimate total training time.

After training for each task, the instructor rated the controls used to perform the task on clarity, effectiveness, efficiency, and simplicity; always in that order. Clarity was defined as the degree to which interface controls were clear and understandable. Effectiveness was defined as the degree to which interface controls allowed the task to be performed. This criterion provided an opportunity for instructors to recommend additional functionality to improve the effectiveness of controls. Efficiency was defined as the degree to which the controls used in performing the task allowed quick performance. Simplicity was defined as the degree to which the logic of using the controls was complex or easy to understand. Extremes of the rating scale for simplicity were anchored with verbal anchors, "Extremely High Simplicity" and "Extremely High Complexity". Rating scales for clarity, effectiveness, and efficiency were like the one described in Figure 3 below except the applicable factor was inserted.

0 = "Do not know"
1 = "No clarity "
2 = "Extremely Low clarity"
3 = "Very Low clarity"
4 = "Below Average clarity"
5 = "Average clarity"
6 = "Above Average clarity"
7 = "High clarity"
8 = "Very High clarity"
9 = "Extremely High clarity".

Figure 3. The rating scale

The process of obtaining ratings guided the instructor to think about specific criteria for usability and provided indicators of order relationships among tasks for each criterion. After rating the control interface for a task on one of the criteria, instructors were asked to identify problems and recommend improvements. At the conclusion of the session, instructors were asked to make summary evaluations of the utility of the training device.

RESULTS

A total of 10 personnel from the 96th Security Forces Squadron took part in the preliminary usability evaluation. This report describes evaluation outcomes from 7 of the participants. These 7 participants are instructors assigned to the 96th GCTS. Of the remaining participants, one was not an instructor and 2 were instructors. However, the approach was changed after instructor 7 to reduce the length of time required for the evaluation. Each of the first 7 instructors dedicated 6 hours to the evaluation including lunch and breaks. The length of time for the original version of the evaluation was regarded too long; so the evaluation period for each instructor was reduced to 4 hours. The 2 instructors who provided input for the 4-hour evaluation will be combined with input from instructors at another training site where the 4-hour version of the evaluation will be applied. The outcomes from the second evaluation will be described in an independent report.

The average age of the 7 instructors was 25.14 years. All were at the journeyman skill level or higher within the security forces career field. Most of the instructors possessed the rank of staff sergeants (71%) and the remainder possessed the rank of Senior Airmen. They had an average of 5.44 years of service and 9.14 months in the position of instructor. They indicated they spend an average of 46 hours per week using the computer and in the preceding year played computer games, like "Ghost Recon™", an average of 8.57 times. They dedicated approximately 42 hours to the evaluation.

Figure 4. Training time by task averaged over instructors
(Summation over tasks = 1.66 hours)

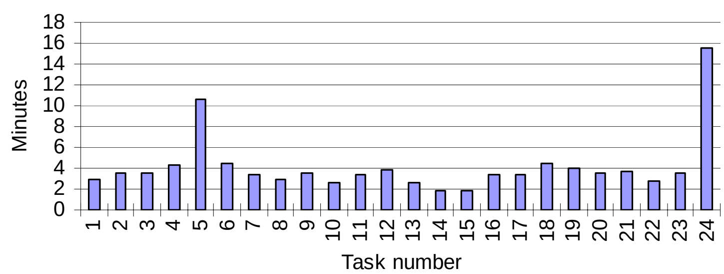
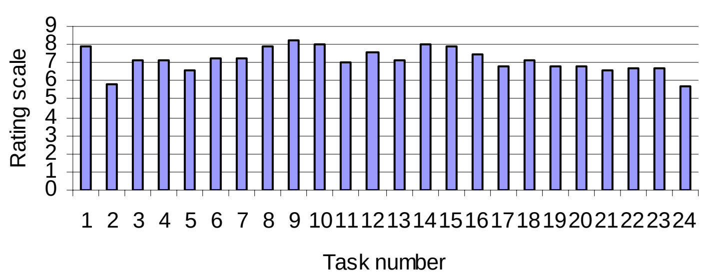


Figure 4 presents the observed training time for tasks averaged over the seven instructors. The training time standard for usability for instructors is 2 hours. The total, observed, average training time for the 24 tasks is 1.66 hours. After training on all tasks, each instructor was asked to estimate the maximum amount of time that could be allocated for instructors to be trained to use the training device. Estimates of the maximum training time for instructors represent the ceiling or upper-bound threshold beyond which additional training time could not be tolerated. The minimum threshold was 2 hours, the maximum was 8 hours, and the median was 3 hours.

Figure 5. Interface clarity averaged over instructors
(Averaged over tasks = 7.15)



For each task, the procedure was to learn how to use controls then rate the interface on clarity, effectiveness, efficiency, and simplicity. Figure 5 presents the average clarity rating for the controls used to perform each of the 24 tasks. The 0-point on the rating scale represents "Do not know" and a rating value of 5 represents "Average Clarity". The aggregate clarity rating for all instructors for each task was above average. These results indicate instructors believed, the interface was easy to understand.

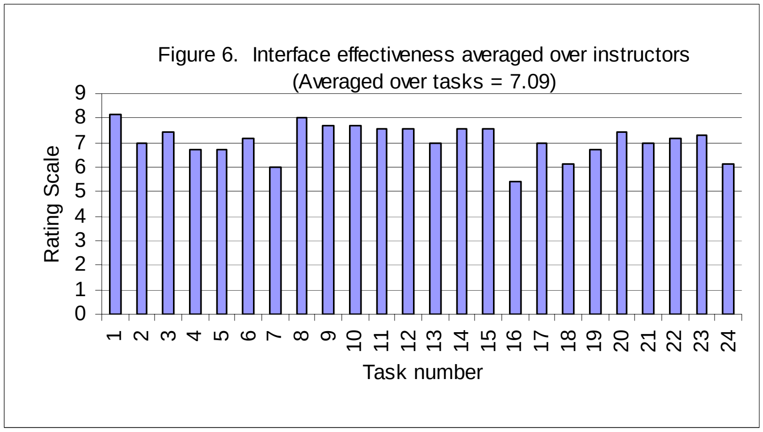


Figure 6 presents the average effectiveness rating for the controls used to perform each of the 24 tasks. The 0-point on the rating scale represents "Do not know" and a rating value of 5 represents "Average Effectiveness". The aggregate effectiveness rating for all tasks was above average.

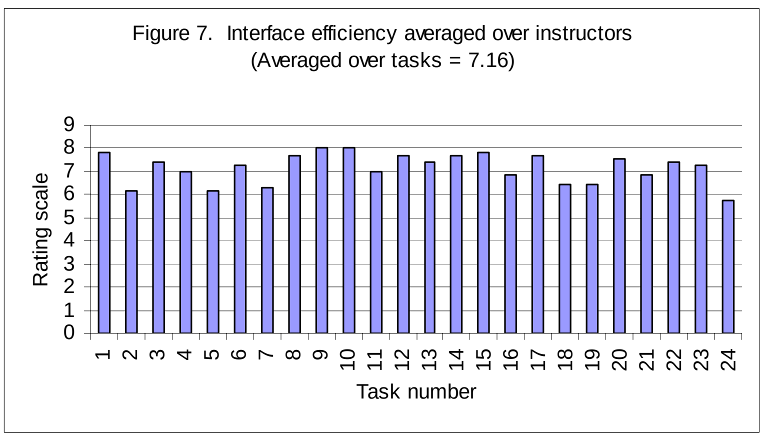


Figure 7 presents the average efficiency rating for the controls used to perform each of the 24 tasks. The 0-point on the rating scale represented "Do not know" and a rating value of 5 represents "Average Efficiency". The aggregate efficiency rating for all tasks was above average. This result indicates instructors believed, the interface allowed them to accomplish tasks quickly. An important design issue became apparent during evaluation of Task 10 (Create, select, move, orient, and delete a sensor; accomplished by using the right click menu). The computer model for the tripwire sensor provides a tripwire 240 meters in length. Several instructors were quick to point out that the model was wrong; 15 meters would be the maximum length for a trip wire. However, one instructor noted that if the tripwire sensor were modeled accurately, 16 sensors would have to be created, located, and oriented to cover the same line covered by the incorrectly modeled, 240 meter tripwire. He recommended the 240 meter tripwire be maintained to improve usability. For computer-based, training simulations, compromises in model accuracy may yield greater usability. Other examples of this point emerged from the preliminary evaluation. Originally, efforts were made to model CGFs only with combat loads of ammunition specific to their weapon so trainees would have to remember to re-supply them when ammunition was expended. However, it became apparent that this was not required for **threat** CGFs. The amount of ammunition carried by threat CGFs need not be modeled accurately. Threat CGFs are an instructors' tool for supporting teaching points. Rather than having an accurate model and forcing instructors to take the time to re-supply ammunition to threat CGFs, greater usability is achieved by modeling threat CGFs with infinite ammunition. In this context, it may be desirable to trade off model accuracy for usability if the intended application is not compromised. The training device is intended for command and control training; so usability is regarded as more important than accurate modeling in some cases.

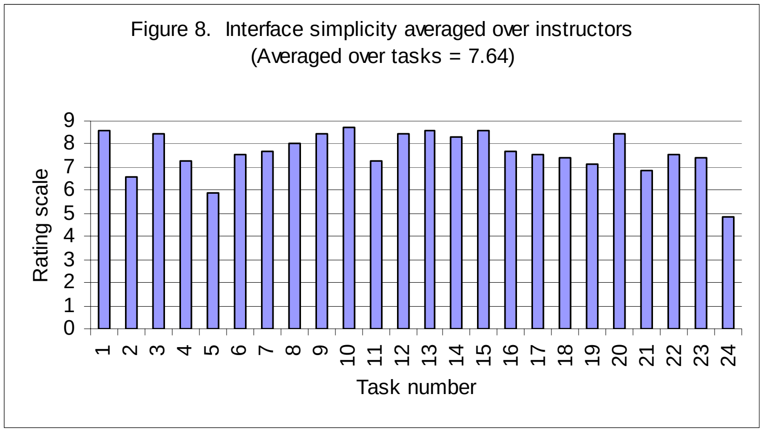


Figure 8. Interface simplicity averaged over instructors (Averaged over tasks = 7.64)

Figure 8 presents the average simplicity rating for the controls used to perform each of the 24 tasks. The 0-point on the rating scale represented "Do not know" and a rating value of 5 represents "Average Simplicity". The aggregate simplicity rating for each task was above average. This result indicates instructors believed the logic of operations required for using the controls was easy to understand.

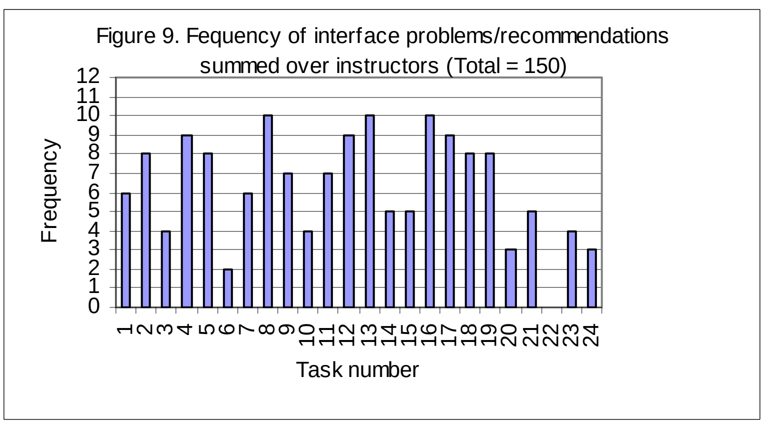


Figure 9. Fequency of interface problems/recommendations summed over instructors (Total = 150)

Figure 9 presents the frequency of interface problems/recommendations summed over instructors. The 7 instructors provided a total of 150 change recommendations. These change recommendations were immediately consolidated, prioritized, and shared with development engineers as guides for system modifications. At the conclusion of the evaluation period, each instructor was asked to make summary evaluations of the value of the training device. Six out of seven instructors indicated the device would add value to training. All seven indicated it could be used to support learning objectives related to defense planning and six out of seven indicated it could support training for decisionmaking and team coordination underlying execution of the defense. Although one instructor did not believe the capability would add value; most instructors believed it would add value and one expressed the desire to have the capability available in the classroom by the end of the year.

DISCUSSION

Although training times obtained during the preliminary evaluation provide useful baseline estimates, these estimates have limitations. Even though observed training time was less than the standard of 2 hours, a few instructors, 3 out of 7, expressed the opinion that the standard time of 2 hours was unrealistically short. There is support for the minority opinion not related to the 24 – task subset. The preliminary usability evaluation did not include the simulation logger or the radio microphone or headset. Training time for these devices will be estimated when the complete training system is evaluated. If tasks for these devices had been included, training time would have been greater.

Total observed training time for the simulation interface underestimates the training time for all the control tasks/options available on the interface delivered with Version 2.0 because only a representative subset of 24 tasks was evaluated. It was necessary for reasons of practicality to limit the number of tasks evaluated. However, knowledge of all control options is not necessary for practical use. How many users of Microsoft Word know how to apply all control options?

There were confounds on training time measurement. Total observed training time for the 24-task sample is an overestimate. During the training period for each task, instructors often discussed problems and recommended needed capabilities. The result was that total observed training time includes time required for training and time for non-training interactions.

Better estimates of observed training time are needed. It would be desirable to obtain a more accurate estimate of observed training time that excludes time for non-training interactions, includes tasks for use of the radio headset, and simulation logger, and a greater number of representative tasks while maintaining a reasonable time for the evaluation period. Multiple evaluation periods may be necessary if acceptable to participants.

The estimate of threshold or maximum training time may have organizational bias. The median threshold was 3 hours. This estimate should be considered in the context of the training production requirement at the 96th GCTS. If instructors at another site with a greater training production requirement had been asked the same question, their estimates could have been less. This suggests estimates of threshold training time should be obtained from instructors at other training sites.

Notwithstanding limitations in estimates of training time, it is important to understand the potential represented by this capability. If training time for this training simulation device were doubled due to including more tasks, it would still be less than the training time required for other simulation capabilities available today. For example, consider the Joint Combined and Tactical Simulation (JCATS). On a separate occasion, the senior author observed security forces personnel being trained on JCATS to support installation security evaluations. Observation of "user training' for JCATS indicated that 7 security forces personnel varying in rank from Staff Sergeant to Master Sergeant satisfactorily learned to task CGFs to move, dismounted and mounted, and shoot, direct and indirect-fire weapons, in 1 and ½ eight-hour training days. Expert controllers indicated they had learned to operate JCATS over a period of months. If the training time for JCATS is the comparison point, the simulation capability emerging from this R&D program could offer a tremendous advantage by requiring much shorter training times for users.

CONCLUSIONS

There were several recommendations that emerged from this effort for structuring usability evaluations of the human-machine interface. Early and frequent evaluation of the interface by end users is recommended. The length of time for the evaluation period should be carefully considered. The 6-hour period initially used in this evaluation was found to be too long and was reduced to 4 hours including time for breaks. If more time is needed to evaluate a greater number of tasks, distributed evaluation periods should be considered. Rather than attempting to obtain evaluations from participants in mass, a one-on-one approach is recommended to minimize the impact on the participating organization and to avoid suppression of recommendations by some evaluators. Also, a structured evaluation approach is recommended. Rather than an unstructured approach where you would ask evaluators what they think about system usability, guide evaluators to think about specific usability criteria. To ensure meaningful input, recommend you train them to use the system before they make their assessments. Finally and most importantly, apply evaluators' input to develop a usable control interface.

REFERENCES

Armstrong, S.D. & Steinberg, R.K. (2002) *Usability Testing.* In Charlton, S.G., O'Brien, T.G. (Eds.) H*andbook of Human Factors Testing and Evaluation* (Chapter 18, pp 403-432). Mahwah, NJ: Lawrence Erlbaum Associates.

Weeks, J. L., Garza, J. S., Archuleta, M. A. & McDonald, L. B. (2001) USAF Security Forces Training Needs. *Proceedings of the 2001 Interservice / Industry Training, Simulation and Education Conference*, Orlando FL.

Weeks, J. L. & McDonald, L. B. (2002). USAF Security Forces Distributed Mission Training: Evolution of the design concept. In, *Proceedings of Spring 2002 Simulation Interoperability Workhop*, Orlando, FL.

APPENDIX

INSTRUCTOR TASK LIST (K_i=24)

Task module A: Control simulation window

Task 1. Use the Select, Next, and Previous buttons to center the map, zoom in 4 levels, zoom out 2 levels, and pan right; accomplished by using the tool bar.

Task 2. Determine distance represented by side of a map grid square by interpreting map coordinates; accomplished by observing map coordinates presented in the simulation window.

Task 3. Create, select, move, and orient a range and bearing line and determine distance represented by side of map grid square and delete line; accomplished by using the right click menu.

Task 4. Turn on / off weapons range fan and field of fire and set field of fire; accomplished by using the Menu bar/ map control menu.

Task 5. Create a terrain box, create an observation point, determine intervisibility from the observation point, delete the terrain box and observation point; accomplished by using the right Click menu.

Task 6. Turn on contour line display and determine elevation with cursor and Altitude window; accomplished by using the Menu bar/ Map control menu.

Task 7. Turn on / off visual display of friendly, threat and neutral entity symbols. Place icon for "Suspected Threat location" on map and do same. Place facility on map and do the same; accomplished by using the Menu bar/ map control menu, and the right click menu.

Task module B: Resource placements

Task 8. Create, select, move, and delete a facility and an aircraft; accomplished by using the right click menu.

Task 9. Create, select, move, and delete an obstacle; accomplished by using the right click menu.

Task 10. Create, select, move, orient, and delete a sensor; accomplished by using the right click menu.

Task 11. Create, select, move, orient, and delete an observation post / listening post; accomplished by using the right click menu.

Task 12. Create, select, move, orient, and delete a friendly computer-generated force (CGF) armed with an M9 and a friendly CGF armed with an M2; accomplished by using the right click menu.

Task 13. Create, select, move, orient, and delete a threat CGF armed with an AK47 and a threat vehicle (BMP-2 AFV); accomplished by using the right click menu.

Task 14. Create, select, move, orient, and delete an armed neutral CGF and an unarmed neutral CGF; accomplished by using the right click menu.

Task 15. Create, select, move, orient, and delete an entry control point (ECP); accomplished by using the right click menu.

Task 16. Place friendly CGF in ECP, turn on primary field of fire and effective range and set primary field of fire; accomplished by using the right click menu.

Task 17. Withdraw and assign telephones to facilities and set loop designations; accomplished by using the Menu bar/tactical planning menu.

Task 18. Withdraw and assign radios to entities, set call signs, and receive and transmit frequencies; accomplished by using the Menu bar/tactical planning menu.

Task 19. Enable audible radio communications and send a location report; accomplished by using the right click menu.

Task 20. Create, save, and send a situation report; accomplished by using the right click menu.

Task Module C: Control Forces

Task 21. Create a route and write a task plan for a HMMWV to patrol a route at maximum speed; accomplished by using the right click menu.

Task 22. Create two waypoints and write a task plan for a CGF to patrol between two waypoints; accomplished by using the right click menu.

Task 23. Write a task plan for one CGF to follow another CGF with rules of engagement set to "Fire at will"; accomplished by using the right click menu.

Task 24. Create a phase line adjacent to two waypoints and write a task plan for a CGF to move to the first waypoint and after crossing the phase line move to the second waypoint; accomplished by using the right click menu.