



AFRL-RH-WP-TR-2007-0109

**AWACS Dialogue Training System
(DTS) Evaluation**

**Michael W. Haas
Human Effectiveness Directorate
Warfighter Interface Division
Wright-Patterson AFB OH 45433**

**Michael Salyer
Northrop Grumman Information Technology
2555 University Blvd.
Fairborn OH 45324**

**Lawrence J. Hettinger
Northrop Grumman Information Technology
7575 Colshire Drive
McLean VA 22102**

August 2007

Final Report for August 2006 to June 2007

**Approved for public release;
distribution is unlimited.**

**Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Interface Division
System Control Interfaces Branch
Wright-Patterson AFB OH 45433**

NOTICE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th Air Base Wing Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

TECHNICAL REVIEW AND APPROVAL

AFRL-RH-WP-TR-2007-0109

**THIS TECHNICAL REPORT HAS BEEN REVIEWED AND IS APPROVED FOR
PUBLICATION.**

FOR THE DIRECTOR

//signed//

Kristen K. Liggett, Ph.D.
Program Manager
System Control Interfaces Branch

//signed//

Daniel G. Goddard
Chief, Warfighter Interface Division
Human Effectiveness Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.					
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 31 Aug 2007		2. REPORT TYPE Final		3. DATES COVERED (From - To) August 2006 – June 2007	
4. TITLE AND SUBTITLE AWACS Dialogue Training System (DTS) Evaluation				5a. CONTRACT NUMBER FA8650-05-D-6633	
				5b. GRANT NUMBER 	
				5c. PROGRAM ELEMENT NUMBER 62202F	
6. AUTHOR(S) Michael W. Haas* Michael Salyer** Lawrence J. Hettinger***				5d. PROJECT NUMBER 7184	
				5e. TASK NUMBER 09	
				5f. WORK UNIT NUMBER T6	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Northrop Grumman Information Technology** 2555 University Blvd Fairborn OH 45324 Northrop Grumman Information Technology*** 7575 Colshire Drive McLean VA 22102				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Materiel Command* Air Force Research Laboratory Human Effectiveness Directorate Warfighter Interfaces Division System Control Interfaces Branch Wright-Patterson AFB OH 45433				10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/RHCI	
				11. SPONSORING/MONITORING AGENCY REPORT NUMBER AFRL-RH-WP-TR-2007-0109	
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES 88 th ABW/PA cleared 09 October 2007, WPAFB-07-0031.					
14. ABSTRACT The objective of this task was to evaluate the features of the AWACS Dialogue Training System. The intent of this system is to reduce the number of personnel resources currently devoted to supporting AWACS WD training. A series of questionnaires were developed to collect information from participants. The results of this evaluation are intended to help inform AETC about the maturity and readiness of the AWACS-Dialogue Training System to support WD training, and provide suggestions for potential improvements.					
15. SUBJECT TERMS AWACS training, dialogue training system, voice recognition					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT SAR	18. NUMBER OF PAGES 64	19a. NAME OF RESPONSIBLE PERSON Kristen K. Liggett
a. REPORT UNCL	b. ABSTRACT UNCL	c. THIS PAGE UNCL			19b. TELEPHONE NUMBER (Include area code)

THIS PAGE LEFT INTENTIONALLY BLANK

TABLE OF CONTENTS

1. INTRODUCTION.....	1
1.1 BACKGROUND.....	1
2. FACILITY DESCRIPTION.....	2
2.1 AWACS DTS	2
2.1.1 MISSION SCENARIOS.....	3
2.2 INTEGRATED EVALUATION FACILITY	5
2.2.1 STATION 1.VBMS.....	5
2.2.2 STATION 2. VBMS RADIO	7
2.2.3 STATION 3. AWACS-DTS	7
2.2.4 STATION 4. JSAF	7
2.2.5 STATION 5. DATA LOGGER.....	8
3. STUDY DESCRIPTION.....	8
3.1 SCOPE	8
3.2 DESIGN	9
3.3 PARTICIPANTS	10
3.4 PROCEDURE.....	10
4. RESULTS.....	11
4.1 QUESTIONNAIRE DATA.....	11
4.1.1 MISSION QUESTIONNAIRE RESULTS.....	11
4.1.2 SUMMARY QUESTIONNAIRE RESULTS	13
4.1.3 FEATURES NEEDING IMPROVEMENT RESULTS	14
4.2 SUBJECTIVE WORKLOAD	15
4.3 FACILITATOR OBSERVATIONS	16
5. DISCUSSION	17
5.1 MISSION QUESTIONNAIRE DATA	17
5.2 SUMMARY QUESTIONNAIRE DATA.....	17
5.3 IMPROVEMENT NEEDS DATA	18
5.4 RESPONSES TO OPEN-ENDED QUESTIONS	19
5.5 DISCUSSION OF PARTICIPANT'S GENERAL IMPRESSIONS.....	21
5.6. SWAT DISCUSSION	21
6. CONCLUSIONS	22
7. RECOMMENDATIONS	23
REFERENCES	24
APPENDIX A. LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS.....	25
APPENDIX B. REFERENCE SHEET FOR VBMS CONTROL SEQUENCES.....	26
APPENDIX C. EVALUATION SCRIPT	27
APPENDIX D. QUESTIONNAIRES	38
APPENDIX E. CONSENT FORM	51
APPENDIX F. SUBJECTIVE WORKLOAD ANALYSIS TECHNIQUE.....	54
APPENDIX G. SUMMARY OF PARTICIPANT RESPONSES	55

LIST OF FIGURES

Figure 1.	AWACS Aircraft	1
Figure 2.	The AWACS-DTS Components	2
Figure 3.	Standard Military Headset.....	3
Figure 4.	Mission reas W-151A and W-151C	4
Figure 5.	The Integrated Evaluation Facility	5
Figure 6.	Bank of AWACS Weapons Director Consoles	6
Figure 7.	VBMS Screen Shot during Engagement	7
Figure 8.	Conversational Interactivity Diagram	8
Figure 9.	Participant Responses for Each Mission.....	12
Figure 10.	Average of Responses to Summary Questionnaires	13
Figure 11.	Responses for Features Needing Improvement.....	14
Figure 12.	Subjective Workload Ratings	15

LIST OF TABLES

Table 1.	Evaluation Script Outline	10
Table 2.	Comparison of Means for Mission Questionnaires	12
Table 3.	Comparison of Means for Summary Questionnaires	13
Table 4.	Comparison of Means for Improvement Features	14
Table 5.	Subjective Workload Analysis of Variance.....	15
Table 6.	Areas of Enhancement Identified by Participants	20

EXECUTIVE SUMMARY

The Air Education and Training Command, AETC/A5TT, requested an evaluation of their Airborne Warning and Control System Dialogue Training System (AWACS-DTS). The evaluation focused on assessing the voice recognition and synthesis capability of the DTS that was intended to simulate the behaviors and voice interactions of fighter pilots training for air-to-air combat and thus, facilitate efficient training for AWACS weapon directors. The DTS was delivered to AFRL/RHCP and integrated into the SAFIRE architecture to begin the evaluation. A series of questionnaires were developed collecting demographic and experience information, as well as subjective ratings and inputs on the effectiveness and utility of the overall training, ease of learning, and quality of voice recognition. Data were collected from highly experienced operators during their use of the DTS. Results of the evaluation indicated that the operators perceived the simulated behaviors and radio calls from the fighters and Eglin Mission Control as being realistic. However, the operators were frustrated by the lack of feedback provided by the DTS and believed they had little, or no, influence over the behavior of the fighter aircraft. Additionally, the operators perceived that the cadence of communication was not dynamic and that this characteristic significantly hindered their interactivity with the DTS. The operators understood how a system such as the DTS could be utilized to potentially improve the training of future weapon directors. However, they believed that the current configuration of the DTS was not ready for use in a formal training setting. The operators also stated that DTS-type technology would also be useful “in-the-field” for trained operators who were deployed and not getting sufficient air combat control practice.

THIS PAGE LEFT INTENTIONALLY BLANK

1. INTRODUCTION

1.1 Background

One of the many roles of the Air Education and Training Command is to develop and evaluate technology that may increase the quality and/or efficiency of training within the Air Force. AETC/A5TT is using their Education and Training Technology Application Program (ETTAP) as a contracting vehicle to acquire advanced training technology. Specifically related to this effort, AETC developed a technology to increase the efficiency of training Weapons Directors for the Airborne Warning and Control System (AWACS) aircraft, shown in Figure 1, by simulating the behaviors of fighter aircraft and voice communications between the AWACS, the air traffic control system, and fighter aircraft typically utilized in air-to-air training missions flown over the Gulf of Mexico.



Figure 1. AWACS Aircraft

In this evaluation, this technology is referred to by the term AWACS Dialog Training System (AWACS-DTS) or as the DTS. This technology was initially integrated by AETC into the AWACS Modeling and Simulation System located at the 325th Air Combat Squadron. However, this capability could not be adequately evaluated in this configuration. Simultaneously to the AETC technology development, AFRL/RHCP was developing and empirically evaluating crew-system interfaces for Air Force systems operated in network-centric environments (Haas, 2006). AFRL/RHCP was interested in the AETC technology for its ability to support system-of-system level human-in-the-loop experiments. AETC requested, and funded, AFRL/RHCP's support in evaluating AWACS-DTS within AFRL facilities.

In preparation for this evaluation, AETC contracted with Southwest Research Institute (SwRI), the developers of the training technology, to re-package the technology into a "stand-alone" system that could be integrated by AFRL/RHCP into its existing simulation architecture. SwRI re-packaged their capability into a single Linux-based micro-processor system that communicated using IEEE Std 1278 Distributed Interactive Simulation (DIS) techniques. The evaluation was

initiated with the integration of the re-packaged stand-alone system, the AWACS-DTS, into AFRL/RHCP's SAFIRE simulation architecture at Wright-Patterson Air Force Base, Ohio.

2. FACILITY DESCRIPTION

2.1 AWACS-DTS

In this evaluation, the AWACS-DTS was treated as a “black-box”. In other words, the details of the hardware and software implementation were not known by the investigators and only the input-output characteristics were observable. The AWACS Dialogue Training System (DTS) consists of a single rack-mountable unit, a David Clark headset, and SWRI-built headset-computer interface system. A block diagram of the AWACS-DTS depicting its relationship to the AMS Training system is shown in Figure 2.

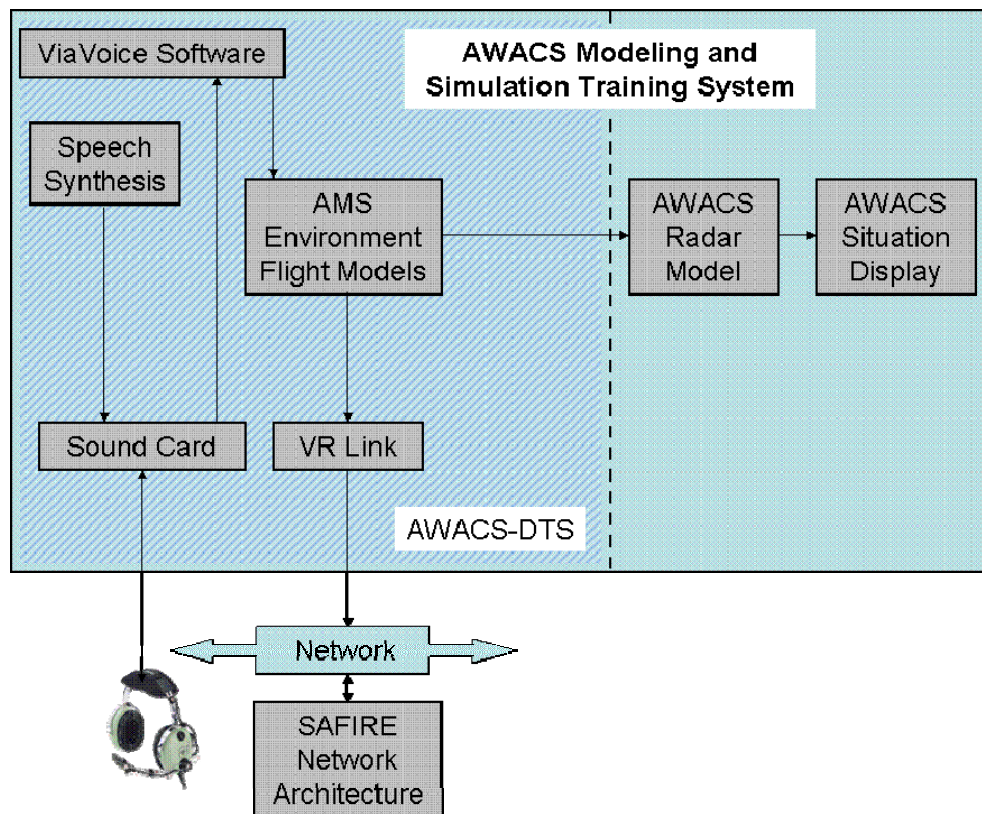


Figure 2. The AWACS-DTS components of the AMS Training system

The AWACS-DTS interacts with a single user via the David Clark microphone and headset-mounted speakers. A similar headset is shown in Figure 3.



Figure 3. Standard Military Headset

The AWACS-DTS provides a speech recognition module that allows the participant to interact with the internally-modeled fighters. The speech recognition module supports a limited vocabulary of standard weapons control radio terminology. The AWACS-DTS also generates Distributed Interactive Simulation (DIS) protocol IEEE-STD-1278 entity state data packets. The AWACS-DTS does not receive DIS data packets and thus can not react to externally generated entities or events. The AWACS-DTS operates at the SECRET level. Two specific air combat training scenarios are implemented in the AWACS-DTS. Detail descriptions of these are contained in Appendix B as TI-3 and ACT-1. The mission scenarios are summarized in the following paragraph.

2.1.1 Mission Scenarios

The mission scenarios simulated are peace-time training scenarios typically flown over the Gulf of Mexico by a flight of four fighter aircraft practicing air combat tactics.

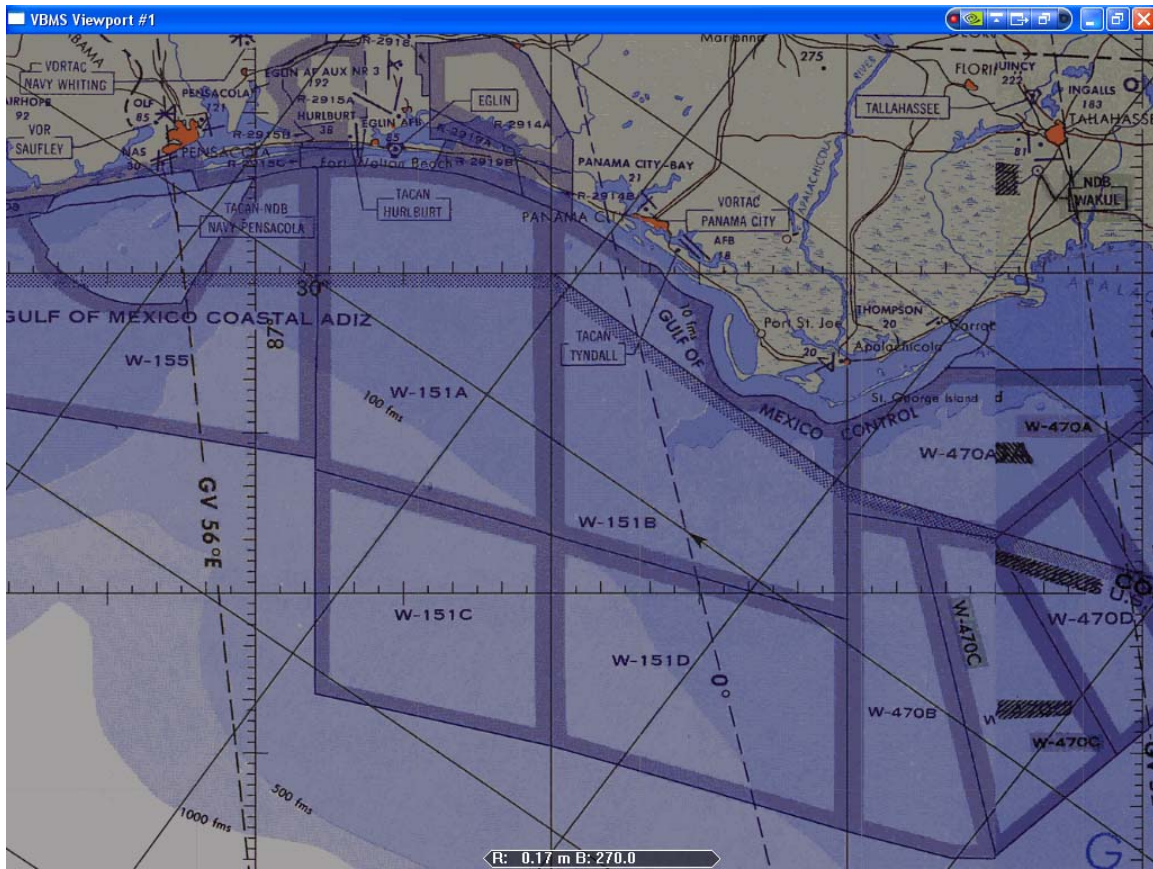


Figure 4. Areas W-151A and W-151C were used in the study

The two mission scenarios used in the evaluation are the Tactical Intercept Level 3 mission (TI-3) and the Air Combat Training Level 1 (ACT-1) mission. Many variations of the two mission scenarios are possible to specify in the AWACS-DTS. Variations are achieved by specifying items such as training area, voices for the fighter aircraft, call signs, IFF codes, etc. In this evaluation, a single set of these items were utilized for each of the participants. Both scenarios commenced with the four fighter aircraft south of the Florida panhandle heading west as if they had recently departed Tyndall AFB. The voice interaction with the weapon director is initiated with a requested handoff of fighter control from Eglin Mission Control to the AWACS. The fighter aircraft continue west and enter the Northeast corner of area WD101A. At this point they perform several equipment checks. Following the checks, two aircraft begin to fly Southwest while the remaining two aircraft orbit a point close to the Northeast corner of area W-151A. A map depicting the location of area W-151A is shown in Figure 4. After several minutes of flying Southwest, the southern two aircraft reverse course and the tactical engagement between the four aircraft begins. Once the engagement is complete, the aircraft again separate and repeat the profile. After three such engagements, the fighters indicate to the AWACS that they are ready to leave

the controlled airspace and return to their airbase. At this point in the mission, the AWACS operator contacts Eglin Mission Control and hands-off control. Both the TI-3 and ACT-1 missions follow a similar pattern with the difference being in the behaviors of the fighter aircraft during the engagements.

2.2 Integrated Evaluation Facility

The AWACS-DTS was integrated into a set of computational resources enabling the participant weapon director to control the visualization the output of the DIS information generated by the AWACS-DTS while verbally communicating with the simulated ground controllers and fighter aircraft. The facility utilized four computer workstations in addition to the AWACS-DTS. All five workstations were networked together and was a mixed Linux/Windows environment. . The five computer workstations are described in the following paragraphs. A block diagram of the integrated evaluation facility is shown as Figure 5.

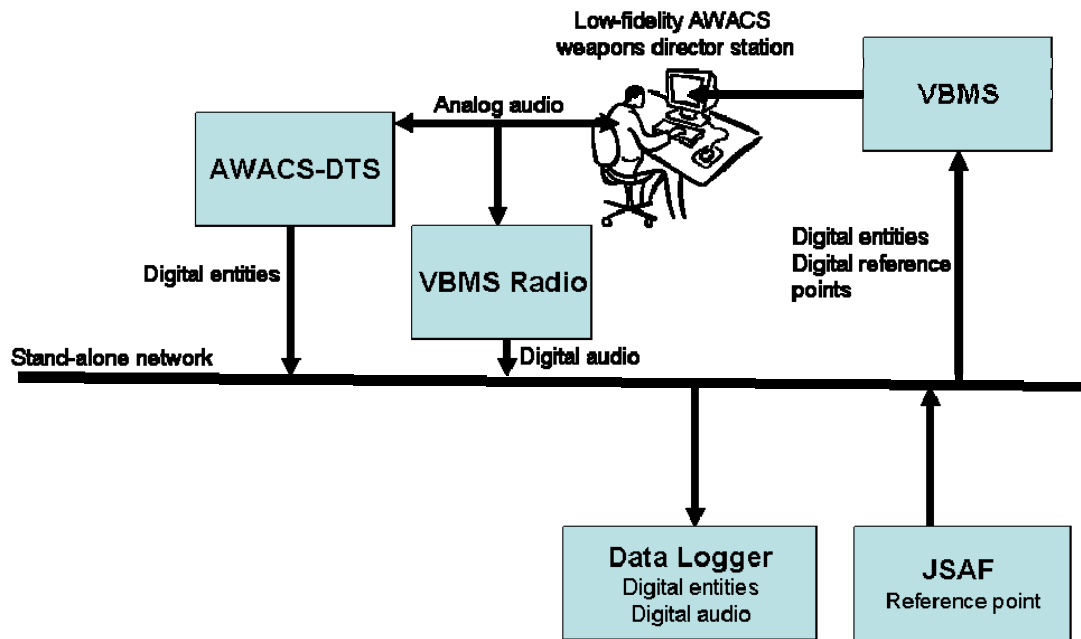


Figure 5. Block diagram of the Integrated Evaluation Facility

2.2.1 Station 1: VBMS. This station served as a low-fidelity simulation of the AWACS weapons director console. A bank of three AWACS weapons director's consoles is shown in the photograph labeled as Figure 6.



Figure 6. AWACS Weapons Director's consoles

The VBMS station was a typical computer workstation with a flat panel computer screen, keyboard, and mouse resting on a desktop. The participant was seated. The VBMS station provided the participant the ability to control the visualization of the control airspace. The control implementation does not replicate those controls on-board the AWACS aircraft however, the visualization, available information, and the functionality of the controls replicates the functionality and information available on-board the aircraft. A “reference sheet” was provided to remind the participants of the keystrokes and mouse clicks needed for actions such as zooming into a visual area of interest or measuring the bearing, range, and altitude of an aircraft relative to a reference point. This “reference sheet” is contained in Appendix B. The participant also used a foot switch as a push-to-talk switch as is used on-board the AWACS aircraft. The software running on this station was the Virtual Battlespace Management System (VBMS) developed by ASC/FI and General Dynamics AIS. A screen shot of the VBMS display similar to that shown during the study is shown in Figure 7. The bearing and range line is depicted as the line running from the yellow reference point to the blue fighter aircraft with the bearing and range shown in numeric form in the

bottom center of the screenshot. Also in this shot, aircraft information is shown in the lower left corner which was activated by rolling the mouse pointer over the aircraft from which information is requested.

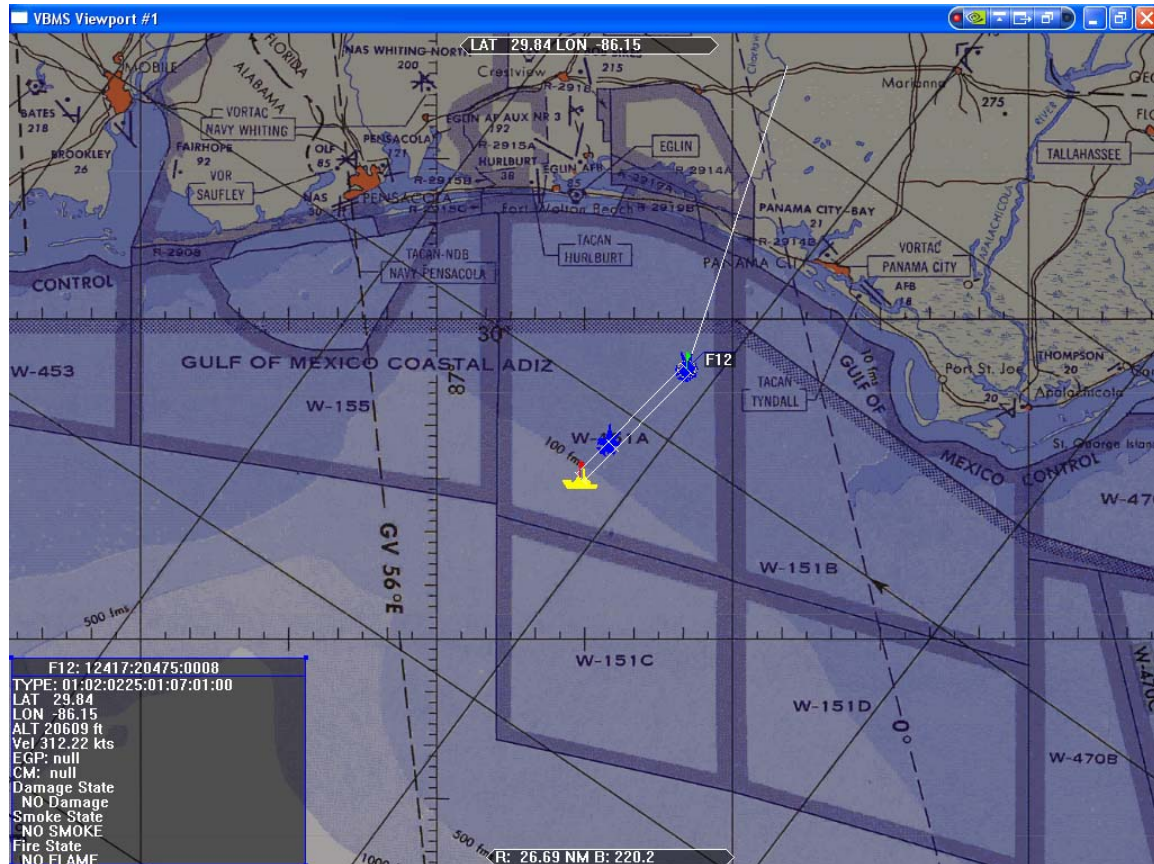


Figure 7. Screen Shot of VBMS during a fighter separation

2.2.2 Station 2. VBMS Radio. This software running on this workstation translated the voice synthesis outputs of the AWACS-DTS and the verbal utterances of the participants into DIS packets transmitted on the network. The software utilized was ModIOS Voice developed by General Dynamics. The networked DIS packets were then available to be recorded in synchrony with the entity state DIS packets. The participant did not interact with this station.

2.2.3 Station 3. AWACS-DTS. The AWACS-DTS was the primary simulation engine within the facility. The participant interacted directly with this station during voice recognition training but not during any other portion of the evaluation.

2.2.4 Station 4. JSAF. This workstation fed the operator's visual display with a precisely placed "permanent" bull's-eye for the participant's use as a reference point for fighter interaction. The participant did not interact with this station.

2.2.5 Station 5. Data Logger. This workstation recorded all DIS data packets on the network during simulation execution. The DIS data packets included the verbal dialogue between the participant and the system as well as the position and attitude of the fighter aircraft. The participant did not interact with this station.

3. STUDY DESCRIPTION

3.1 Scope

A simple diagram of the time course of a conversation between two sources is seen in Figure 8. In Figure 8, State A could represent the weapon director, State B the AWACS-DTS, State M a pause in conversation, and State D a “stepped-on” transmission, in this case the weapon director “stepping-on” the AWACS-DTS. A regular pattern including State A, State M, and State B is sometimes called a dialogue cadence and would be considered a “normal” speech pattern in which information flows between the two sources normally. When State M

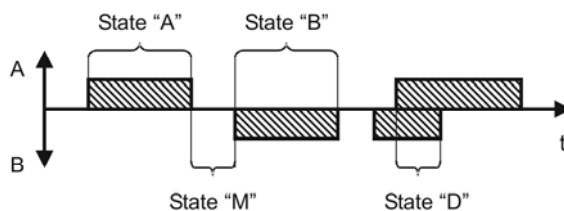


Figure 8. Conversational Interactivity Diagram

becomes too short or too long, or State D begins to occur, the dialogue becomes labored and information flow becomes restricted leading to frustration if one, or both, of the sources is a human. Hammer et al (Hammer, 2005) describes the dialogue cadence pattern as “conversational interactivity”. He describes several causes for State D, such as natural interruptions, which he terms active interruptions, or a time delays in a transmission line, causing passive interruptions. Hammer also describes the affect of time delay on conversational interactivity and a metric for measuring the affect of active and passive interruptions on conversational interactivity (Hammer, 2005).

The AWACS-DTS is distinguished by its use of computer-based voice recognition and synthesis capabilities in combination with the embedded behaviors of the fighter aircraft under control of the weapon director. These capabilities were the focus of the evaluation. The AWACS-DTS, in its present configuration, is primarily a technology demonstration tool. This evaluation attempts to fully identify the strengths and weaknesses of the technology within the AWACS-DTS as it can be observed by the use of the system. The AWACS-DTS could be thought of as a computer-based training system and indeed does possess some characteristics typical of a computer-based training system. However, because it was not developed to be a stand-alone computer-based

training system, this evaluation does not pursue a methodology that would completely evaluate a computer-based training system.

Specifically, this evaluation focused on an assessment of the system's voice recognition and synthesis capability as a means of supporting effective "trainer" - trainee interactions. The evaluation methods used were intended to assess the quality of the verbal interaction between the trainee and automated "trainer" (i.e., AWACS-DTS) on the assumption that the quality of the training afforded by the use of this system will be a direct function of the quality of the verbal interaction.

No attempt was made in this evaluation to objectively assess the overall training effectiveness of the AWACS-DTS; however, subjective measures of potential training effectiveness were collected from crew members who have been trained, and in some cases trainers, using other methods and offer opinions from a perspective of highly experienced weapons directors.

3.2 Design

This evaluation was designed in a manner that allowed each participant to be exposed to the two scenarios (TI-3 and ACT-1) while operating the AWACS-DTS. A series of questionnaires were developed to extract information from the participants on their qualitative evaluation of the training system. These questionnaires are shown in Appendix D. The questionnaires include a demographic questionnaire, mission questionnaire (one for each mission type evaluated, and summary evaluation questionnaire. The demographics questionnaire was developed to collect background and experience information on the participants. The mission questionnaire was developed to individually assess the TI-3 and ACT-1 missions. The questionnaires collected quantitative ratings (using a seven point Likert scale where 1=Disagree, 4= No Opinion, and 7=Agree) and qualitative information on the participants' perceptions about the mission training (e.g. ease of learning, dialogue success, effectiveness of training, etc.). Participants were also asked to assess whether they felt their verbal interactions and dialogue with AWACS-DTS system were successful (i.e. were communications clear, timely, and consistently appropriate given the situation and time in which they occurred during each mission). A summary evaluation questionnaire was developed and used to collect participant perspective and rating information on the AWACS-DTS system as a whole including the quality of the voice recognition component, overall potential training effectiveness, and general assessment of the AWACS-DTS – particularly, the opportunity for interaction with the system, and the quality of feedback from the AWACS-DTS about the state of the system and student progress (in terms of success or failure).

In addition to completing the questionnaires, at the end of each session, each participant was interviewed by the facilitator to discuss and record general observations on the training. Participants were also given the opportunity to make additional evaluative comments.

3.3 Participants

A human use protocol was submitted for review to the AFRL/Wright Site Institutional Review Board (IRB). The IRB granted approval for the evaluation in April 07. The participant consent form contained in the protocol is shown in Appendix E.

Seven operators with extensive WD experience were recruited to serve as participants in the AWACS-DTS evaluation. The participants had a total of 91 years of experience as WDs in systems including AWACS (primarily), Navy E-2, etc. The least experienced operator had seven years. The breadth of participant experience included positions as a NATO surveillance officer, NATO instructor, crew commander, senior AWACS WD/instructor, E-2C battle manager/air controller/mission commander, weapons control officer, battle staff duty officer, Airborne Communications Center (ABCC) Strike Controller, and ground radar system controller.

3.4 Procedure

An evaluation script, shown in Appendix C, was developed to in-brief and out-brief each participant on their role in the evaluation. The evaluation script ensured consistency in the execution of the evaluation by identifying all the steps required during the evaluation and described the instructions given to each participant.

The steps described in the evaluation script are summarized in Table 1:

Step Number	Step Description
1	Equipment Setup and Check-out
2	Participant Introduction
3	Execute Consent Form
4	Participant Pre-brief
5	Demographic Questionnaire
6	SWAT Introduction and Card Sort
7	Training the Voice Recognition System
8	Participant Training on VBMS use
9	TI-3 and ACT-1 Mission Familiarization
10	Data collection and Questionnaires
11	Evaluation De-briefing

Table 1. Evaluation Script Outline

Response data from participants was collected as described in the evaluation script. Three questionnaires were developed to gather demographic information,

subjective ratings, and written responses from participants were administered to each participant as described in the evaluation script. The questionnaires are shown in Appendix D.

Participants were tested individually in single 3-4 hour test sessions. Prior to testing, each participant completed and signed the consent form and were briefed on the purpose of the evaluation as well as their role in it. The Subjective Workload Assessment Technique (SWAT) was briefed to each participant. As described in Reid, 1989. Each participant trained the speech recognition capability and received familiarization training with the AWACS-DTS and the controller workstation operating VBMS.

Each participant was given opportunities to ask questions regarding the purposes of the evaluation and its specific procedures during familiarization. Following familiarization, each participant served as a weapons director during two data collection trials, the first using TI-3 and second using ACT-1.

During each mission, while the fighters were separating, each participant was asked to give a SWAT score. At the end of each mission, the participant was also asked to give a SWAT score for the same mission flown in the jet.

After each mission, participants completed the appropriate questionnaire, one for the TI-3 mission and one for the ACT-1 mission. At the end of both data collection missions, each participant was asked to complete the summary evaluation questionnaire. Following completion of the questionnaires, each of the mission trials was replayed and discussed with each participant using the AWACS-DTS log file as a cue for mission events. The purpose of replaying each mission was to identify and categorize interactions of interest for the participant. The final de-brief of the participant was then accomplished by the facilitator.

4. RESULTS

4.1 Questionnaire Data

4.1.1 Mission Questionnaire Results. The complete results of the Likert scale ratings provided by each participant for the TI-3 Mission Questionnaire, ACT-1 Mission Questionnaire, and overall Summary Questionnaire are included in Appendix G. The quantitative scores for the mission questionnaires are depicted graphically in Figure 9. A Comparison of Means was performed using Scheffe's procedure to tease apart the questions that elicited "negative" responses and those that elicited "positive" responses. A rejection level of 0.1 was used for the statistical analysis. Negative responses represent areas of concern while positive responses represent characteristics that are implemented especially well. For questions with reversed polarity wording, the polarity of response was reversed for graphical and analytical

consistency. The results of the Comparison of Means is shown below the graphical depiction of the data in Table 2.

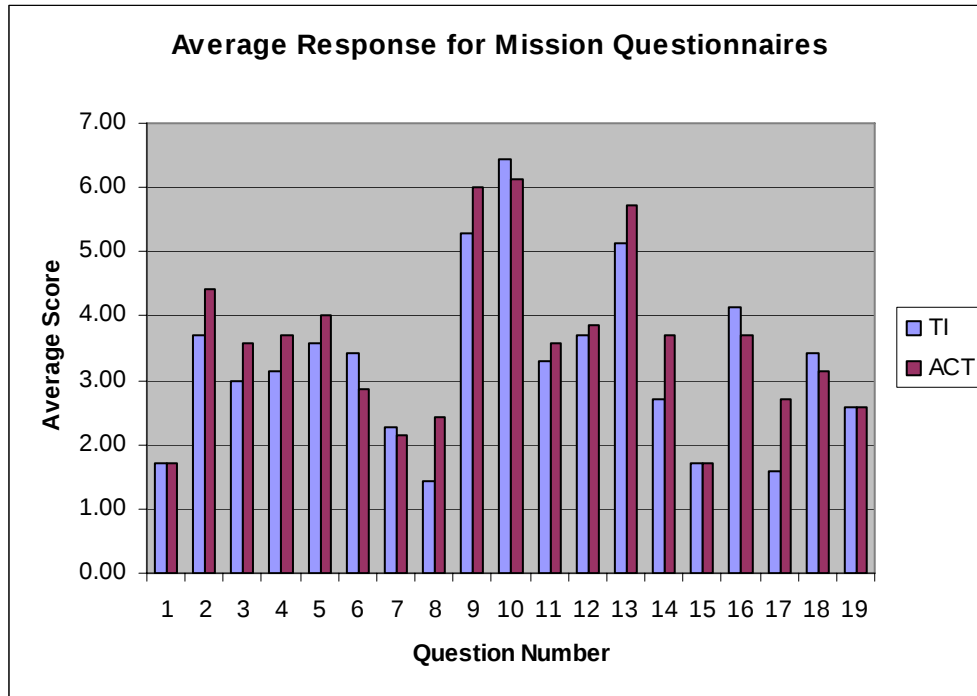


Figure 9. Participant Responses for Each Mission

SCHEFFE COMPARISON OF MEANS OF RESP BY Question		
Question	MEAN	HOMOGENEOUS GROUPS
10	6.2857	I
9	5.6429	II
13	5.4286	III
2	4.0714	IIII
5	3.7857	IIII
12	3.7857	IIII
18	3.7143	IIII
4	3.4286	IIII
11	3.4286	IIII
3	3.2857	IIII
14	3.2143	IIII
6	3.1429	IIII
16	3.0714	IIII
19	2.5714	I
7	2.2143	II
17	2.1429	III
8	1.9286	IV
15	1.7143	V
1	1.7143	V

THERE ARE 4 GROUPS IN WHICH THE MEANS ARE NOT SIGNIFICANTLY DIFFERENT FROM ONE ANOTHER.

CRITICAL F VALUE 1.473 REJECTION LEVEL 0.100
 CRITICAL VALUE FOR COMPARISON 3.2877
 STANDARD ERROR FOR COMPARISON 0.6385
 ERROR TERM USED: RESIDUAL, 247 DF

Table 2. Comparison of Means for Mission Questionnaires

4.1.2 Summary Questionnaire Results. The quantitative scores for the summary questionnaires are depicted graphically in Figure 10. A Comparison of Means was performed using Scheffe's procedure to tease apart the questions that elicited "negative" responses and those that elicited "positive" responses. A rejection level of 0.1 was used for the statistical analysis. Negative responses represent areas of concern while positive responses represent characteristics that are implemented especially well. For questions with reversed polarity wording, the polarity of response was reversed for graphical and analytical consistency. The results of the Comparison of Means is shown following the graphical depiction of the data in Table 3.

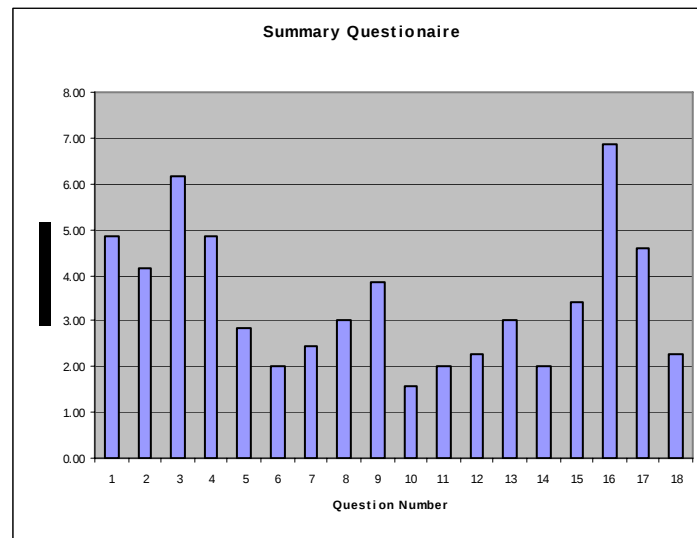


Figure 10. Average of Responses to Summary Questionnaire

SCHEFFE COMPARISON OF MEANS OF RESP BY QUESTION		
Question	MEAN	HOMOGENEOUS GROUPS
D2	6.8571	I
A1	6.2857	II
A3	6.1429	II
B3	6.1429	II
A4	5.8571	III
A2	5.5714	III
B1	4.8571	III
B4	4.8571	III
D3	4.5714	III
B2	4.1429	III
C2	3.8571	III
D1	3.4286	III
C3	3.0000	III
C1	3.0000	III
B5	2.8571	III
B7	2.4286	III
C5	2.2857	III
D4	2.2857	III
D6	2.0000	III
C4	2.0000	III
C7	2.0000	III
C3	1.5714	III

THERE ARE 3 GROUPS IN WHICH THE MEANS ARE NOT SIGNIFICANTLY DIFFERENT FROM ONE ANOTHER.

CRITICAL F VALUE 1.467 REJECTION LEVEL 0.100
 CRITICAL VALUE FOR COMPARISON 4.5650
 STANDARD ERROR FOR COMPARISON 0.8225
 ERROR TERM USED: RESIDUAL, 132 DF

Table 3. Comparison of Means for Summary Questionnaires

4.1.3 Features Needing Improvement Results. The quantitative scores for the features needing improvement questionns are depicted graphically in Figure . A Comparison of Means was performed using Scheffe's procedure to tease apart the questions that elicited "negative" responses and those that elicited "positive" responses. A rejection level of 0.1 was used for the statistical analysis. Negative responses represent areas of concern while positive responses represent characteristics that are implemented especially well. For questions with reversed polarity wording, the polarity of response was reversed for graphical and analytical consistency. The results of the Comparison of Means is shown following the graphical depiction of the data in Table 4.

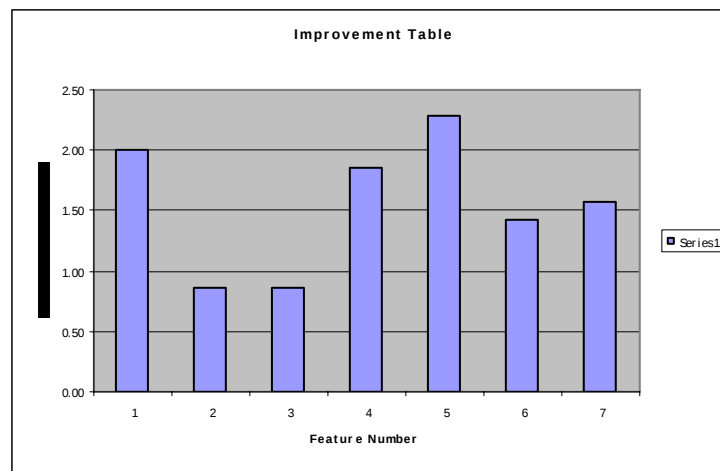


Figure 11. Responses for Features Needing Improvement

SCHEFFE COMPARISON OF MEANS OF RESP BY QUESTION		
Question	MEAN	HOMOGENEOUS GROUPS
F5	2.6667	I
F1	2.3333	II
F4	2.1667	II
F7	1.8333	II
F6	1.6667	II
F2	1.0000	..I
F3	1.0000	..I
THERE ARE 2 GROUPS IN WHICH THE MEANS ARE NOT SIGNIFICANTLY DIFFERENT FROM ONE ANOTHER.		
CRITICAL F VALUE 1.950 REJECTION LEVEL 0.100		
CRITICAL VALUE FOR COMPARISON 1.4029		
STANDARD ERROR FOR COMPARISON 0.4102		
ERROR TERM USED: RESIDUAL 35 DF		

Table 4. Comparison of Means for Improvement Features

4.2 Subjective Workload

The workload of the operators was measured using the Subjective Workload Assessment Technique (SWAT) (Appendix F) described by Reid et al. (Reid, 1989). The three-digit SWAT ratings taken after each engagement within a mission were converted into percentages (Reid, 1989) and then averaged for each operator resulting in a single average SWAT rating for the evaluation session for that operator. In addition, the three-digit SWAT ratings taken after each mission based on the operator's memories of training missions in which they had served as air warfare officers, were converted into percentages and then averaged. This resulted in a single SWAT rating for each operator based on their "real-world" experience during training missions of the type flown during the evaluation.

The two sets of SWAT ratings, one set based on the evaluation session, and the second set based on "real-world" experiences, are graphically depicted in Figure 12. Subjective Workload Ratings. The environment is indicated as either "RW", for the "real-world" environment, or "Sim", for the simulated environment. The error bars depict standard error. The two sets of SWAT rating were compared using a one-way ANOVA and the results are shown in Figure 12 and in Table 5.

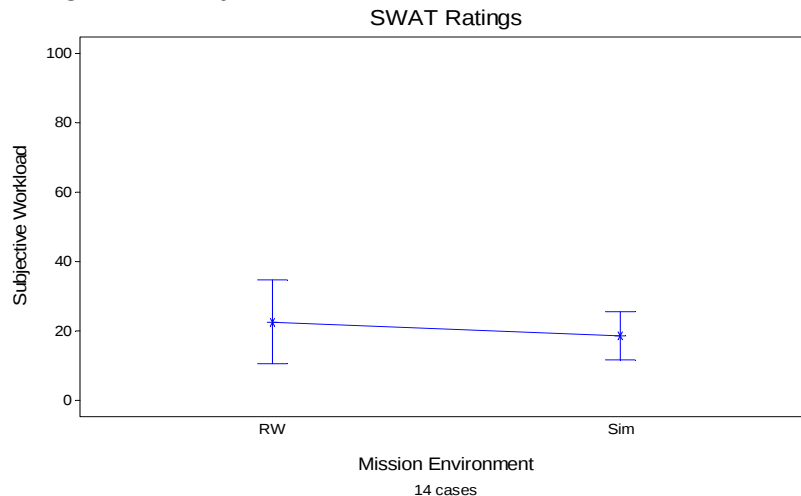


Figure 12. Subjective Workload Ratings

ANALYSIS OF VARIANCE TABLE FOR WORKLOAD					
SOURCE	DF	SS	MS	F	P
OPERATOR (A)	6	7005.89	1167.65		
RATING_TY (B)	1	56.0000	56.0000	0.29	0.6096
A*B	6	1158.47	193.078		
TOTAL	13	8220.36			

Table 5. Subjective Workload Analysis of Variance

4.3 Facilitator Observations

The following comments are based on the observations and general impressions of the facilitator responsible for conducting each of the AWACS-DTS evaluations:

Lack of feedback, lack of cadence -- The biggest problem with the system appeared to be the lack of feedback. Whether something was done correctly or incorrectly the participants received no information. Many of the participants assumed being ignored or having no response from the fighters was a positive message. The lack of feedback severely hampered the ability of the students to establish a cadence with the fighter pilots. The constant back and forth confirmation and acknowledgement that is part of the communications between fighter and operator could not be established.

Observing not participating -- The operator was more of an observer than a participant. The experience seemed to be little better than watching a training video. As an example, all fourteen trails went autonomous, meaning that during the FAA handoff, the highly experienced operators did not use wording that the AWACS-DTS would accept. The average time to go autonomous during the TI-3 mission was 8.29 minutes and 8.43 minutes during the ACT mission.

Negative training -- One concern is that this system will confuse and frustrate new and inexperienced operators. Judging from the amount of frustration from experienced operators, who are confident of their ability to make the correct calls, an inexperienced operator might even be negatively influenced by this system.

Limited vocabulary -- Many of the calls the participants make received no response from the fighters or the fighters responded with "say again". It was not apparent whether the fighter did not understand the words or the meaning. The system seemed to have a very limited vocabulary which seemed to contribute to the frustration level of the participants.

Voice pitch cutoff -- The system appeared to have difficulties with higher pitched voices during the voice recognition training. Females and males with higher pitched voices appeared to have more difficulty in making themselves understood. Voice recognition sessions took longer to complete because of unrecognized words and did not appear to improve as the training progressed. The inability to handle high-pitched, and possibly low-pitched, voices limits the usefulness of the system by limiting the students who can successfully use the voice recognition. A restraint on high pitched voices also presents a problem in high-stress situations during which the pitch of the student's voice may rise and therefore be out of range.

False perceptions of success -- Many of the successful communications were actually just part of the pre-programmed actions of the system. For example, the aircraft checking, directing the aircraft to return to the airspace, or calling "fighters on" were all activities the aircraft did on all missions or intercepts whether the controller said anything or not. The participants, however, were under the impression that the fighter had finally understood them or listened to them. This

could possibly lead to reinforcing incorrect calls from students who thought they were finally saying the right call.

Voice recognition training --The voice recognition training lasted from a minimum of 10 minutes to a maximum of 37 minutes with an average of 19.17 minutes.

5. DISCUSSION

5.1 Mission Questionnaire Data

An interpretation of the comparison of means analysis indicates that questions 1, 7, 8, 15, and 17 elicited primarily negative responses. These questions are listed below.

1. I completed the TI-3 training session and aircraft did not go autonomous.
7. I felt I was able to establish and maintain a consistent rhythm of Dialogue with the AWACS-DTS system during the TI-3 mission.
8. I felt that the fighter element was reacting appropriately to my communications.
15. Eglin Mission Control and the fighters never interrupted my communication.
17. I had appropriate control over the fighters.

A low score on question 1 was expected since the fighters announced that they were going autonomous on every trial. Questions 7 and 15 reflect the inability to establish a cadence with the radio transmissions. Questions 8 and 17 are related. A lack of appropriate reactions from the fighters gave the participant a feeling of having no control over the fighters and generated frustration.

An interpretation of the comparison of means analysis also indicates that questions 9 and 10 elicited primarily positive responses. These questions are listed below.

▪ 9. I used the map display to effectively control and advise the fighters.
▪ 10. The appropriate level of feedback was provided by the facilitator before and after the mission.

The positive response from question 9 indicates the operators had little problem adapting from the airborne control station they use operationally and the VBMS system used in this evaluation.

5.2 Summary Questionnaire Data

An interpretation of the comparison of means analysis indicates that questions B6, C3, C4, C5, C7, and D4 elicited primarily negative responses. These

questions are listed below. Question D4 has a negative polarity and its response was reversed for this evaluation indicating significant transmissions were stepped on during the evaluation.

B6. I found the system to be flexible.
C3.The system provided feedback to me on my progress throughout the mission.
C4. The feedback I received from the system was always appropriate to the situation and my performance.
C5. The feedback I received from the system was consistent throughout the mission.
C7. The AWACS-DTS training was as effective as traditional training received for a Weapons Directors (WD).
D4. My radio transmissions were stepped-on only a few times..

The negative responses to these questions indicates the operators were not getting enough information back from the AWACS-DTS to aid them in reducing the number of “stepped-on” .transmissions. Question C7 indicates the operators did not believe the AWACS-DTS should be used in a training environment as a stand-alone tool.

An interpretation of the comparison of means analysis also indicates that questions A1, A3, B3, and D2 elicited primarily positive responses. These questions are listed below.

▪ A1. The voice recognition system was easy to learn.
▪ A3. The map display was easy to learn
▪ B3. The map display was easy to use.
▪ D2. “Natural” voice interaction is important to promote effective training

These positive responses indicate that overall, the operators did not have trouble training the AWACS-DTS voice model and that the use of the VBMS was relatively intuitive. The response to question D2 indicates the operators felt that the dialogue properties elicited by the AWACS-DTS would be important in its ability to train future weapon directors.

5.3 Improvement Needs Data

An interpretation of the comparison of means analysis indicates that question F5 elicited primarily negative responses indicating a need for improvement in this area. This question is listed below.

F5. Eglin Mission Control System

An interpretation of the comparison of means analysis also indicates that questions F2 and F3 elicited primarily positive responses. These questions are listed below.

▪ F2. Audio sound quality
▪ F3. Control Stations (Map Viewer)

The participants found the audio portrayal of the simulated fighters and mission control to be highly realistic. They were also able to easily utilize the VBMS control station.

5.4 Responses to open-ended questions

The questionnaires and the debriefing by the facilitator gave the participants opportunities to express both positive and negative perceptions of the technology, and system operation. These expressions are captured in the following discussion.

The participants had several suggestions for improving the system. The most prominent is to improve the voice recognition. Improvements would include a larger vocabulary for the voice recognition and the inclusion of recognizable feedback. Both of these modifications would allow the student to feel comfortable that he or she was heard and understood. Controllers express their commands a little differently but with standardized buzzwords. The system has to recognize those words within the controller input and respond/react appropriately. Otherwise, the student will become frustrated. In lieu of these modifications, using human operators to produce the Dialogue would also be welcome. Human instructors would also have the benefit of providing experienced advice and feedback.

Feedback, or the lack of it, was very important to the participants. Controllers need meaningful feedback when calls are understood whether it be audio (wilco, roger, copy) or visual. Visual feedback could consist of text message (separate Dialogue box) for what the system heard and/or what the system was looking for or whether the student successfully responded. Feedback on erroneous transmissions could be displayed after two wrong attempts. Summary performance feedback at the end of each mission would also be helpful. This information could include the percentage of calls understood.

The lack of feedback is closely connected with the perception of the quality of the voice recognition capability. Participants believed that the Dialogue was too limited and the simulation required structured (“canned”) responses. There was a lot of frustration due to unacknowledged transmissions and repeated “say again” from the fighters and Eglin Mission. The lack of response from the aircraft of

commands or confirmation of directives also disturbed the participants. The lack of feedback and confirmations prevented the controller from establishing a cadence with the fighters and contributed to the feeling of no control over the fighters. Participants felt more like passive observers than active participants, especially when the fighters did not follow commands. The fighters were going to run the entire mission no matter what the controller said.

It was unclear whether the lack of flexibility was from the inadequacies of the voice recognition or due to misunderstood objectives of the training. The participants disliked the inability to converse/start/commit a mission or to KIO when necessary. There was no response to KIO calls when fuel state was at Bingo and to KIO calls for safety (busted airspace several times).

The participants identified various areas of the AWACS-DTS where they thought improvements were needed. Table 6 highlights some of the comments provided by participants.

Feature	Improvements Needed
Voice recognition	- Has limited functionality, forcing student to repeat transmissions - Some words are difficult to establish - If the aircraft are not ready– the student should be able to ask questions trying to ascertain status - Need some sort of feedback
Fighter action	- Fighter symbology flipped back and forth confusing the true heading (seemed to be turned wrong direction based on target calls) - More positive responses from the fighters - Communication between controller and pilots is not realistic - Controller has little or no effect on outcome (another threat, vector/snap somewhere else)
Eglin mission control system	- Did not respond appropriately to any radio transmissions - Inability to hand off with direction - Lack of positive response from ground control agencies
Voice recognition training	- Increase/Include vocabulary – all Military Classification Manual (MCM) 3-1 terms - It needs better training so it will work – because it seemed not to work
Voice recognition usage	- Include vocabulary – all MCM 3-1 terms - Good in theory, not so good in reality - Immediate correction of mistakes (e.g. “spike range 20” computer text would show “spike range 10”)

Table 6. Areas of Enhancement Identified by Participants

5.5 Discussion of Participant's General Impressions

The operators saw the potential of having the ability to self train without the stress of a real-life session, as well as the fact that the AWACS-DTS did not require other personnel resources (e.g. enlisted personnel performing as fighter pilots), possibly resulting in significant savings to the Government. The participants liked the potential capability of controlling the fighters and the mission control (software) tools to control the fighters. For the most part, the controllers thought the simulations appeared realistic. The fighter intercepts, fighter calls of bandit locations, sequence of radar handoff, and timing for the scenario were realistic. However, the setup time was for each intercept was too long, adding no value, and the "busted airspace" is usually not so blatant. While there appeared to be no confusion over the identification of the planes, some participants felt that the picture requests came at odd times, and there was some annoyance regarding the directional symbology of the fighters.

Participants evaluating the system expected that they could speak as a WD and the system would understand responses/commands by the controller and respond appropriately. They also expected to be able to influence the actions of the fighters by providing information and instructions via voice commands. However, there is no readily identifiable positive feedback provided by the system that indicates student radio transmissions were received and understood by the system (i.e. that communications transmitted by the student influence the actions of the fighters during the mission). If the system goes autonomous (the fighters say "Going Autonomous"), the student can still engage the fighters and receive a response. The fighters respond to the student transmissions with "Say Again" or with silence. A "Say Again" response means that the fighter did not receive a coherent transmission via the voice recognition or the system could not recognize the command. Silence from the fighters means the transmission was not received or the student's transmission was understood or the fighters ignored the transmission. The definition for successfully completing a training session with the AWACS-DTS was unknown. Therefore, for this evaluation, it was decided that a successful mission would mean a successful handoff from and to the simulated FAA operator without the fighters going autonomous.

5.6 SWAT Discussion

Based on the ANOVA results in Table 5, the null hypothesis can not be rejected indicating the two sets of SWAT ratings are not significantly different. This result strongly suggests that the workload of the operators created by controlling the 4 F-15s in the two mission types evaluated during this study is similar to the workload experienced by these operators when controlling similar missions in the "real-world".

Based on the results of the questionnaires, participants agreed that the requests for the SWAT score during the mission intercepts did not disrupt the mission activities.

6. CONCLUSIONS

The participants were clearly frustrated by the lack of feedback in the system. Several types of feedback were needed. The only feedback apparent were the almost constant “say again” from the fighters and Eglin Mission. The lack of any other active feedback created an implied feedback from the silence of the fighters. However, participants could not be sure what silences or “say again” meant. Several options could be assumed from the fighters’ silence and the requests to repeat the transmission:

- Controller responded appropriately and was understood,
- Controller responded inappropriately but was understood,
- Controller responded but the command/buzzwords were not understood,
- Controller responded but the transmission/translation was not clear,
- Controller responded appropriately but the system was simulating pilot error, or
- Controller responded appropriately but the simulation failed.

This lack gave the participants the perception that the system was not communicating with them and that the controllers had no influence on the fighters. The affirmative responses that are present in human-to-human communication create a cadence between the controller and fighter giving both a sense of working as a team. Without the affirmative responses, inexperienced controllers may develop inappropriate habits/expectations and experienced controllers will be frustrated. The lack of response from the fighters created a perception of being a passive observer rather than a participant. The system has to incorporate more feedback – audio and visual – to meet the minimal expectations of a CBT. Participants believe training needs to reflect the way they fight in real life or at the least “real-life” training.

The system failed to respond appropriately in several types of situations. These may be due to a failure in the simulation or in the voice recognition. Fighters did not respond to KIO calls or corrections when they were exceeding the boundaries of the designated airspace or when fuel levels were low. The fighters took an exceedingly long time to set up the intercepts leaving the participant with no value added during this wait time.

The actions and radio calls from the fighters and Eglin Mission were very realistic according to the participants. The participants were also very pleased at the possibility of the cost savings and the ability to solo train. However, they all agree that the system is not ready in its current state. The voice recognition appears to work poorly. Most transmissions had to be repeated multiple times. Adding to the frustration was the frequency with which the participants’ transmissions were interrupted by the fighters and by Eglin Mission. The source of these errors is unknown but it is speculated that system processing time impacts the development of proper dialogue timing between the system and the user. A

cycle can develop in which the user attempts to repeat a communication when the system is not able to totally accept it resulting in more processing time and the production of a “say again” while the user is repeating their last communication.

Training the voice model was sometimes very difficult to complete. There was no way to skip difficult words. Some words had to be repeated and in one case, had to be repeated more than 50 times before the participant could continue. In some cases, as the training continued, the model did not seem to “learn” or recognize previously established words. The training also seemed to have greater difficulty with higher pitched voices.

7. RECOMMENDATIONS

It is clear from this evaluation that the AWACS-DTS, as it is currently implemented, should not be included in a training environment. The lack of flexibility in the dialogue model, coupled with the lack of feedback to the user, greatly reduces the usability of the system and potentially greatly reduces its ability to support independent training of operators. In its current state and used without the aid of an instructor, the AWACS-DTS may be detrimental to training in that a student, instead of focusing on control aspects, could become too frustrated or distracted with the system’s characteristics. If used in conjunction with an instructor’s assistance, the AWACS-DTS may be of some limited use in a training curriculum. If development of the technology is to proceed, it is clear that the voice recognition components need to be greatly improved especially in the area of supporting flexible verbal cadences. The vocabulary of keywords needs to be expanded to accommodate variations in commands and dialogue in general. This flexibility should be extensively tested in the acceptance testing phase of the system using experienced operators. There was some anecdotal evidence suggesting female voices were more difficult to train in the system than male voices. The range of voices may need to accommodate a larger range of human voice characteristics.

There is great potential to enhance student training with a significantly enhanced AWACS-DTS capability. Arguably, the most important improvement needed for the AWACS-DTS is a more extensive system of feedback. The feedback should include both audio and textual information. Failures and successes need to be addressed and acknowledged. Simple feedback to the user for radio calls that were received, understood, and appropriate would make a significant difference in the usability of the system. Transmissions that are not fully received, misunderstood or inappropriate need to be addressed as well.

There are potentially additional opportunities for use of the technologies represented by the AWACS-DTS. Training for first-time controllers, supplemental and continual training for active controllers, practice for

recertification, training for flights involving UAVs, and area familiarization before deployment. These are listed below.

- Supplement instructor training, after hours, or as warm-up training for other training
 - Continuation training for experienced controllers in the field
 - Area familiarization before deployment
 - Flights dealing with Unmanned air Vehicles (UAVs)
 - First time controller to see how sortie works
 - Basic/Initial AWACS training
 - Predator operators and TACP
 - BQ training (i.e. new controllers) prior to first flight
 - Any training where players are at different nodes in an NCW exercise
 - Initial WD training, Instructor Qualification Course (IQC), and Mission Qualification Training (MQT)
 - B-1 Weapons Engineering Officer (WEO) training – threat reaction, radar rejoin – aircrew interaction
 - Electronic warfare school

Because of the great number of additional applications for this technology, it is recommended that AETC continue spiral development of the functionality of the AWACS-DTS ensuring the technical causes of the current issues are identified and excluded from future designs. The spiral development should include Air Force operational crew member participation in the development and implementation of test metrics and test methodology involving voice recognition capability and overall usability to include metrics such as Hammer's conversational temperature (Hammer, 2005).

REFERENCES

Reid, G.B., Potter, S.S., and Bressler, J.R. (1989) *Subjective Workload Assessment Technique (SWAT): A User's Guide*, AAMRL-TR-89-023, Harry G. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, OH.

Haas, M.W., Middendorf, M.S. and Ulling, S. (2006) *The SAFIRE and an Initial Network-centric Warfare Evaluation*. (Report No. AFRL-HE-WP-TR-06-0129). Wright-Patterson AFB, OH: Air Force Research Laboratory.

Hammer, F, Reichl, P., and Raake, A. (2005), The Well-Tempered Conversation: Interactivity, Delay and Perceptual VoIP Quality, Proc. IEEE ICC'05, Seoul, Korea, May 2005.

APPENDIX A. LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS

AETC	Air Education and Training Command
ABCC	Airborne Communications Center
ACS	Air Combat Squadron
ACT	Air Combat Training
AFRL	Air Force Research Laboratory
AMS	AWACS Modeling and Simulation System
ANOVA	Analysis Of Variance
AWACS	Airborne Warning and Control System
BQ	Basic Qualification
CBT	Computer-based training
DIS	Distributed Interactive Simulation
DTS	Dialogue Training System
ETTAP	Education and Training Technology Application Program
FAA	Federal Aviation Administration
HEC	Warfighter Interface Division
IQC	Instructor Qualification Course
IRB	Institutional Review Board
KIO	Knock It Off
MCM	Military Classification Manual
MQT	Mission Qualification Training
NATO	North Atlantic Treaty Organisation
SAFIRE	Synthetic and Human Forces in a Research Environment
SWAT	Subjective Workload Assessment Technique
TACP	Tactical Air Control Post
TI	Tactical Intercept
UAV	Unmanned Air Vehicle
VBMS	Virtual Battle Management System
WD	Weapons Director
WEO	Weapons Engineering Officer

APPENDIX B. REFERENCE SHEET FOR VBMS CONTROL SEQUENCES

5/7/2007

VBMS Reference Card

Zoom	Activate zoom	Left Shift + Left Mouse Button
	Stop zoom	Left Shift + Middle Mouse Button
	Clear zoom	Left Shift + Right Mouse Button
LDU	Place object	Right Ctrl + Left Mouse Button
	Move object	Right Ctrl + Middle Mouse Button
	Select object	Right Ctrl + Right Mouse Button
Range Point/Object	Pick starting point	Left Ctrl + Left Mouse Button
	Pick ending point	Left Ctrl + Middle Mouse Button
	Clear selection	Left Ctrl + Right Mouse Button
PVM	Center on location	Left Mouse Button
	Pan NSEW	Middle Mouse Button
Route point	Place route point	Right Shift + Left Mouse Button
	Move route point	Right Shift + Middle Mouse Button
	Select route point	Right Shift + Right Mouse Button

SWAT (1=Easy, 3=Hard):

Time Load

Mental Effort Load

Psychological Stress Load

APPENDIX C. EVALUATION SCRIPT

6/11/2007

ETTAP Evaluation in Synthetic and Human Forces in a Research Environment (SAFIRE)

Evaluation of the AWACS Dialogue Training System (DTS)

Evaluation Script (V2.1)

1 Setup

1.1 For Evaluation

1.2 Set Up Equipment for Each Subject

Plan to arrive 15 minutes prior to the arrival of the subject.

Start each computer and log in:

1.2.1 Station One (Bottom): VBMS

User ID:	xxxx
Password:	xxxx

Double-click VBMS_DIS icon

Double-click VBMS icon

Open Config file AETC.TVTF

Double-click "Washington State" or Click OK

Reposition map to the area

1.2.2 Station One (Top): Vbms radio

User ID:	vbmsradio
Password:	vbmsradio

Double-click VBMS_RADIO

1.2.3 Station Three (Top): Jsafgateway

User ID:	xxxx
Password:	xxxx

Double-click JSAF Run_jsafgate.sh icon.

Set the permanent bull's-eye: File – Load Scenario – Bull's-eye

6/11/2007

1.2.4 Station Three (Bottom): Logger

Double-click **Logger icon**.

User ID:	xxxx
Password:	xxxx

1.2.5 Station Two: SwRI

Setup new account for the subject:

Click **Instructor**

User ID:	xxxx
Password:	xxxx

Click **Add user**

Enter student ID and write down the new name

Click **Exit**

Start Training:

Click **Student**

User ID:	xxxx
Password:	xxxx

Enter the new Student ID

Select **Train Speech Recognition**

(Wait to click OK until the student is ready to begin)

2 Greet Subject outside the vault

- Check subject ID against the name that has a verified clearance (from the email)
- Tell them to leave all electronic devices in vestibule
- Open and enter inner door
- Sign subject into the vault on the Visitor's Log

3 Conduct Pre-Briefing

Thank you for coming today and welcome to AWACS DTS evaluation Lab. Today we are going to be evaluating an AWACS DTS.

The goal of this evaluation is to examine the quality of the interaction between you and the AWACS DTS system.

Since the interaction between you and the system will be primarily vocal, you'll need to complete voice recognition training so that the training system can build a model of your voice patterns. You'll conduct the voice recognition training on this system.

6/11/2007

After voice recognition training, you'll receive training on how to control VBMS.

During the evaluation, you'll be asked to provide a Subjective Workload Assessment Technique (SWAT) score to rate your feelings about the workload you just experienced. I'll provide you with more complete explanation of SWAT and what you'll be doing in a minute.

You'll complete two missions using the system. One mission is a tactical intercept training mission and the other is air combat training. During these missions, you'll be wearing these headphones, using VBMS to watch the events unfold, and using the foot switch and microphone to verbally interact with the FAA and the pilots.

It's very important that you remember that we're evaluating the system and not you. You're the expert. You should interact with the system as you feel appropriate. Say the things that you would normally say in the way that you would normally say them. We're evaluating how well the system responds to you.

I will be watching your interaction with the system and taking notes from time to time. For the most part, you can just ignore me; but at certain points in the mission, I will ask you for a SWAT score. I want you to give me a quick SWAT score and return to the mission.

After each mission, you'll complete a short questionnaire about your interactions with the system during the mission. I'll ask you a few questions about your experience while it's still fresh in your mind.

After the first mission, you'll be given a 10-minute break. You may ask for other breaks as necessary, but we ask that you refrain from taking a break during a mission. Once you start a mission, it cannot be paused. We'd have to start the entire mission over.

After you complete the second mission, you'll complete the short questionnaire and another questionnaire that asks you about the entire experience from the voice recognition to using the VBMS.

Then I'll provide you a short debriefing and give you a chance to ask me any questions you might have about the evaluation.

Do you have any questions now?

4 Complete Consent Form

[\[Provide the Experiment Consent Form\]](#)

6/11/2007

4.1 Purpose

AFRL/HE is asking you to participate in an evaluation of the AWACS DTS. By participating in this evaluation, you will help us understand the quality of the system and your interactions with it.

4.2 Evaluation Environment

The evaluation takes place in the SAFIRE Lab where you will be observed as you use the training system.

4.3 Information Collected

We will record information about how you use the system. We will ask you to fill out questionnaires, and we will interview you. A digital recording will be made of your voice along with the voices of the computer-generated actors in the mission. We will use the information you give us, along with information we collect from other people, to evaluate and recommend ways to improve the training system. We may use your voice and verbal statements, but not your name for the purposes of evaluating the system and showing the results of the evaluation.

4.4 Comfort

We have scheduled one break for you, but you may ask for a break at any time you wish. Inform the evaluation administrator when you would like to take a break.

4.5 Freedom to Withdraw

You may withdraw from this evaluation at any time.

4.6 Freedom to Ask Questions

If you have any questions, you may ask the evaluation administrator now or at any time during the evaluation.

5 Complete Experience/Demographic Questionnaire

[\[Provide the AWASC DTS Demographic Questionnaire.\]](#)

6 Introduction to SWAT (Outer Hallway)

During the course of this evaluation, you will be asked to quantify the mental workload required to complete the tasks you will be performing. Mental workload refers to how hard you work to accomplish some tasks, or an entire job. The workload imposed on you at any point in time consists of various dimensions which contribute to the subjective feeling of workload. For this evaluation, we will use the Subjective Workload Assessment Technique (SWAT) to assess workload based on three dimensions - (1) Time Load, (2) Mental Effort Load, and (3) Psychological Stress Load.

6/11/2007

Let's go over a brief description and example of each of these dimensions.

[\[Provide the subject the dimension descriptions/examples packet and briefly go over it with them\]](#)

Each of the three dimensions we've discussed also have been assigned three levels (least workload to highest workload) respectively assigned as 1, 2, and 3. The dimensions and their levels are described on the cards I will provide you.

[\[Provide the subject the SWAT cards\]](#)

Since there are three dimensions and three levels for each dimension, there are 27 possible combinations, and therefore 27 cards, with each combination represented by a card. One of the most important features of SWAT is its unique scoring system. SWAT uses a procedure to find separate scoring weights for each level of a dimension. Then it determines a distinctive workload scale for each individual or group. In order to develop your individual scale, we need information from you regarding the amount of workload *you feel* is imposed by the various combinations of the dimensions we've talked about – (1) Time Load, (2) Mental Effort Load, and (3) Psychological Stress Load.

I would like for you to sort the cards in order of your perception of increasing workload on a person. In essence, from the *best* case situation or *lowest* in terms of your perception of workload, to the *worst* case situation or *highest* in terms of workload. Let me re-emphasize that your ranking should be based on your *general* perception of workload based on these dimensions, and not any specific task.

You may use any strategy you choose to order the cards. However, one strategy that may be helpful is to arrange the cards into three preliminary stacks representing high, moderate, and low workload. You can then rank order cards within each stack and, if necessary, exchange cards between stacks. Once, your comfortable with the sorting of each stack, you can recombine the stacks, and re-check the ordering of all 27 cards. Again, reordering cards as you see fit. Pay attention to the descriptors on the cards to make your judgments, so you can become comfortable with the levels of the dimensions.

During the evaluation, there will be distinct points in time during the missions that I will ask you to provide a SWAT score based on your

6/11/2007

opinion of the mental workload required to perform the task. "1" representing the lowest load for a particular dimension and "3" the highest workload. The score you will provide will consist of a number from each of the three dimensions. For example, a possible SWAT score is "1-2-2". This represents a score of "1" for Time Load, a "2" for Mental Effort Load, and a "2" for Psychological Stress Load. Again, "1" representing the lowest load for a particular dimension and "3" the highest workload. Do you have any questions? If not let's begin the SWAT card sort.

[\[Perform SWAT Card Sort and Record Sort Results\]](#)

7 Introduction to Evaluation Environment

The AWACS DTS consists of this single computer that generates the simulation and operates at the SECRET level. The system uses speech recognition input to allow you to interact with the simulated fighters. The speech recognition system provides a limited vocabulary of standard weapons director radio terminology. The fighter pilots and FAA radio transmissions use real voices to simulate these actors in the training system. For the voice recognition training, you will be using this workstation.

For the remainder of the evaluation, you will use this workstation. On this workstation, you will find VBMS. You can use VBMS to observe the mission events as they progress. A "reference sheet" is provided to remind you of some of the important commands and shortcuts for interacting with VBMS. You'll also receive some training to get used to VBMS prior to starting the first mission.

The headset is used for both voice recognition training and the missions. In the voice recognition training you do not need to press any buttons or use the foot switch to make it work. During the missions, however, you will need to use the foot switch to talk to the system.

This workstation up here processes the voice information from you to the system and from the system back to you. You won't need to use this workstation at all.

This workstation interacts with VBMS and provides a "permanent" bull's-eye. This workstation will be used to record the vocal dialog between you and the system as well as collect other data about your interaction with the VBMS system.

8 Voice Recognition Training

Let's begin the Voice Recognition Training. You'll use this workstation for this training.

6/11/2007

There are three parts to this training. When you're finished, the system will have a model of your voice that it can use to recognize what you're saying.

As you start the training, say each word individually, starting with A-M-S. As you progress along, you can start to tie words together until you are speaking normally.

As you read each word successfully, it turns grey. If the system didn't recognize what you said, it will turn the word red and underline it. Read this word again and continue as normal. From time to time, the system turns a period or space red and underlined. Simply read the word after it.

If you get stuck on a word, try saying it again in the same way. If that doesn't work try saying the word along with the next word or two.

Are you ready to begin?

Place the headset on and start reading the text on screen when you're ready.

9 VBMS Training

Let's begin the VBMS Training. You'll use this workstation for this training.

There is a "reference sheet" that lists the available commands and how to access them. Take some time to familiarize yourself with it and let me know when you are comfortable with using it.

Select **Start an Exercise (Area: W470A)**

I'll start a mission so that you have something to work with as you get familiar with the functionality of VBMS. This training will be self-guided exploration, but you should at least try to zoom in and out, get range, bearing, and altitude details for each of the objects on screen, and roll over objects to become familiar with the information that is displayed. Identify the aircraft by ID/digit.

You may have 10-15 minutes to get familiar with VBMS. If you need more time, just let me know. When you're comfortable using VBMS we'll begin the first mission.

<Explain the permanent bull's-eye>

6/11/2007

10 TI-3 Mission

10.1 Setup

10.1.1 Final Instructions to Subject

Instruct subject to put on headset and remind him to use the foot switch.

Inform the subject of the Mode 3 and the Frequency.

10.1.2 Mission Setup

Setup and start the AWACS DTS and set up the TI-3 mission.

Use the following parameters to set up this mission:

[\[Provide the Mission 1 Flight Card\]](#)

Verify that the planes are displayed on VBMS.

10.2 Mission Briefing

This is a typical 4-ship 2-v-2 tactical intercept training mission. The FAA will be handing off a group of F-15's, flying out of Eglin, to Hydra, which is you. At the end of the mission you will hand the fighters back to the FAA. The mission area is in W151AC on a frequency 326.3.

All 2-V-2 intercepts will begin with a "fight's on" call from the lead pilot. All intercepts will terminate with a "knock it off" call. The lead pilot is responsible for calling "knock it off" when the desired learning objective is achieved or is no longer available.

During the first intercept, which is a single group intercept, the fighters should practice BVR weapons employment, multi-targeting and achieving tally-two pre-merge. BVR kills may be called, but the bandits will not kill remove to enable multi-targeting in the visual arena. Subsequent intercepts will be against both Azimuth and Range presentations, and will follow Non-RTO Option 2 kill removal procedures.

All Engagements will be limited maneuvering. Bandits that arrive at a training rule stop will terminate and proceed to their point for the next setup. If the fighter element reaches a training rule stop, a termination will be observed, but fighters may continue to flow if it is conducive to DLO attainment.

10.3 Execution

10.3.1 Data Collection

On the data logging workstation:

- a. Click the red record button on each on the two recorder panels. Click the second button as quickly as possible after the first button.
- b. Verify that both panels are recording.

Record observations during the mission. For each observation, record time and event.

6/11/2007

10.3.2 SWAT Scoring

After each intercept, as the planes separate, ask for a SWAT score.

After the third intercept, ask for a SWAT score.

10.3.3 Mission Termination

When the mission is complete:

- a. Stop the AWACS DTS; (Ctrl + Esc)
- b. Ask the subject for projected SWAT score regarding real-life training mission (Swat score in the jet);
- c. Stop the data logging software;
- d. Ask the subject to complete the End of Mission Questionnaire (reading and explaining the directions); and
- e. Rename the logged files.

10.4 Break

10.5 TI-3 End of Mission Questionnaire

[\[Provide the AWACS DTS ACT-1 Mission Questionnaire\]](#)

10.6 Mission Debriefing

<Mission Debriefing questions/interview here.>

11 ACT-1 Mission

11.1 Setup

11.1.1 Final Instructions to Subject

Instruct subject to put on headset and remind him to use the foot switch.

Inform the subject of the Mode 3 and the Frequency.

11.1.2 Mission Setup

Setup and start the AWACS DTS and set up the ACT-3 mission.

Use the following parameters to set up this mission:

[\[Provide the Mission 2 Flight Card\]](#)

Verify that the planes are displayed on VBMS.

11.2 Mission Briefing

This is a typical 4-ship 2-v-2 Air Combat training mission. The FAA will be handing off a group of F-15's, flying out of Eglin, to Hydra, which is you. At the end of the mission you will hand the fighters back to the FAA. The mission area is in W151AC on a frequency 326.3.

6/11/2007

Eagles are tasked to protect a high value target. The overall risk level for this mission will be high. Select a DEZ based on the airspace and the threat. Plan to skate / short skate / launch & decide based on a willingness to go to a merge according to your position in relation to the DEZ.

The fighter mission objective is to protect target for 15-minute VUL period. The bandit mission objective is to attack and destroy blue fighter caps (and bombs on target within VUL period if strikers are available.)

The fight will begin at the beginning of the VUL period. No fights on call will be made. The intent is to continue the engagement to a logical conclusion, including post-merge flow and reset of the CAP or egress to the FRA at the end of the VUL or when fuel or weapons expended dictates. The lead pilot will terminate the engagement when, in his judgment, the DLOs have been met or are no longer attainable.

If there is a premature knock-it-off on the first engagement and there is sufficient fuel and airspace time available for a second engagement, the lead pilot will restart the fight with a "fight's on" and fight for 15 minutes or until DLOs are achieved.

11.3 Execution

11.3.1 Data Collection

On the data logging workstation:

- a. Double-click Logger.
- b. Click the red record button on each on the two recorder panels. Click the second button as quickly as possible after the first button.
- c. Verify that both panels are recording.

Record observations during the mission. For each observation, record time and event.

11.3.2 SWAT Scoring

After each intercept, as the planes separate, ask for a SWAT score.

After the third intercept, ask for a SWAT score.

11.3.3 Mission Termination

When the mission is complete:

- a. Stop the AWACS DTS; (Ctrl + Esc)
- b. Ask the subject for projected SWAT score regarding real-life training mission;
- c. Stop the data logging software;
- d. Ask the subject to complete the End of Mission Questionnaire; and
- e. Rename the logged files.

11.4 Break

11.5 ACT-1 End of Mission Questionnaire

[\[Provide the AWACS DTS TI-3 Mission Questionnaire\]](#)

6/11/2007

11.6 Mission Debriefing

<Mission Debriefing questions/interview here.>

12 Evaluation Questionnaire

[\[Provide the AWASC DTS Summary Evaluation Questionnaire\]](#)

13 Evaluation Debriefing

<Evaluation Debriefing here.>

14 End of Evaluation

When the evaluation is complete, sign out the subject in the Visitor's Log.

Escort the subject to the vestibule to collect any belongings.

Escort the subject through the vault doors.

APPENDIX D. QUESTIONNAIRES

6/11/2007

ETTAP Evaluation in Synthetic and Human Forces in a Research Environment (SAFIRE)

Evaluation of the AWACS Dialogue Training System (DTS)

Subject Demographic Questionnaire

1. Name _____
2. Today's Date _____
3. Branch of Service _____
4. Rank/Grade _____
5. Organization _____
6. Comm. Phone _____
7. Email _____
8. Current AFSC/Duty Title _____
9. Education – Degree(s) Held:

10. Relevant Experience (Current or Past)
 ___ AWACS/Weapons Director/AWO
 Approximate dates _____
 Yrs. Experience _____

 ___ Similar Weapons Director Position/Experience:
 Title _____

6/11/2007

Platform/System _____

Yrs. Experience _____

11. Are you rated as an AWACS Weapons Director? If not, are you rated on any other command control (C2) system (please specify)?

13. Have you ever completed a computer-based training (CBT) or web-based training course? If so, summarize the type of training and what was your general impression of the training in terms of overall effectiveness?

14. Have you ever had any experience with computer-based, voice recognition systems (in training applications or any other application)? If so, please (a) briefly describe and (b) indicate your impression of the quality interaction (e.g., poor, adequate, excellent, etc.).

15. Have you ever used an interactive training simulator? If so, please briefly describe (a) the training you received and approximate dates, and (b) your impression of the overall quality of training you received with each (e.g., poor, adequate, excellent, etc.)

6/11/2007

ETTAP Evaluation in Synthetic and Human Forces in a Research Environment (SAFIRE)

Evaluation of the AWACS Dialogue Training System (DTS)

Tactical Intercept Mission (TI-3) Questionnaire

Instructions: Please indicate your level of agreement or disagreement with the statements below.
1 = Completely Disagree, 4 = No Opinion, 7 = Completely Agree, or Not Applicable (N/A).

	①	②	③	④	⑤	⑥	⑦	N/A ⑧
1. I completed the TI-3 training session <u>and</u> aircraft did not go autonomous.	①	②	③	④	⑤	⑥	⑦	⑧
2. The TI-3 training with the AWACS-DTS was realistic.	①	②	③	④	⑤	⑥	⑦	⑧
3. The TI-3 mission was effective.	①	②	③	④	⑤	⑥	⑦	⑧
4. Verbal communications / responses from the AWACS-DTS during the TI-3 mission were consistently <i>clear</i> ?	①	②	③	④	⑤	⑥	⑦	⑧
5. Verbal communications / responses from the AWACS-DTS during the TI-3 mission were consistently <i>appropriate</i> given the situation and time in which they occurred?	①	②	③	④	⑤	⑥	⑦	⑧
6. Verbal communications / responses from the AWACS-DTS during the TI-3 mission were consistently <i>timely</i> ?	①	②	③	④	⑤	⑥	⑦	⑧
7. I felt I was able to establish and maintain a consistent rhythm of dialogue with the AWACS-DTS system during the TI-3 mission.	①	②	③	④	⑤	⑥	⑦	⑧
8. I felt that the fighter element was reacting appropriately to my communications.	①	②	③	④	⑤	⑥	⑦	⑧
9. I used the map display to effectively control and advise the fighters.	①	②	③	④	⑤	⑥	⑦	⑧
10. The appropriate level of feedback was provided by the facilitator before and after the mission.	①	②	③	④	⑤	⑥	⑦	⑧

6/11/2007

- | | | | | | | | | |
|--|---|---|---|---|---|---|---|---|
| 11. Overall, I was satisfied with the ease of completing my tasks during the mission. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 12. Overall, I was satisfied with the support information (online-line help, error messages, scenario / mission data) provided for the TI-3 mission. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 13. Overall, I am satisfied with the amount of time it took to complete the TI-3 training session. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 14. The fighter elements performed as I expected. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 15. Eglin Mission Control and the fighters never interrupted my communication. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 16. I believe the system wanted me to do or say something that was not appropriate to the mission. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 17. I had appropriate control over the fighters. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 18. Do you believe the behaviors or actions were supported or reinforced that would reduce the effectiveness of Air Weapons Officer. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 19. Proper skills needed to provide command and control to effectively support this mission are reinforced by the AWACS DTS. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 20. What did you <i>like</i> the most about the system mission? | | | | | | | | |
| 21. Do you believe the system would have been useful during your training? | | | | | | | | |
| 22. What kind of typical CBT provided level of support (online help, performance feedback) have you used? Which one of these would you like to have during this mission? | | | | | | | | |

6/11/2007

23. What did you *dislike* the most about the system mission?

24. Was there any particular point(s) in the TI-3 mission where you felt extremely frustrated – more so than at other points in the session? If so, briefly describe this instance(s) and the reason or cause for frustration.

25. Do you have any suggestions for improving the realism of the mission?

26. Did you experience confusion over which call sign or aircraft were simulating the friendly 2-ship?

27. Did requests for the SWAT score during the 2-ship separation disrupt your mission activities?

6/11/2007

ETTAP Evaluation in Synthetic and Human Forces in a Research Environment (SAFIRE)

Evaluation of the AWACS Dialogue Training System (DTS)

Air Combat Mission (ACT-1) Questionnaire

Instructions: Please indicate your level of agreement or disagreement with the statements below.
1 = Completely Disagree, 4 = No Opinion, 7 = Completely Agree, or Not Applicable (N/A).

	①	②	③	④	⑤	⑥	⑦	N/A ⑧
1. I completed the ACT-1 training session <u>and</u> aircraft did not go autonomous.	①	②	③	④	⑤	⑥	⑦	⑧
2. The ACT-1 mission was realistic.	①	②	③	④	⑤	⑥	⑦	⑧
3. The ACT-1 mission was effective.	①	②	③	④	⑤	⑥	⑦	⑧
4. Verbal communications / responses from the AWACS-DTS during the ACT-1 mission were consistently <i>clear</i> ?	①	②	③	④	⑤	⑥	⑦	⑧
5. Verbal communications / responses from the AWACS-DTS during the ACT-1 mission were consistently <i>appropriate</i> given the situation and time in which they occurred?	①	②	③	④	⑤	⑥	⑦	⑧
6. Verbal communications / responses from the AWACS-DTS during the ACT-1 mission were consistently <i>timely</i> ?	①	②	③	④	⑤	⑥	⑦	⑧
7. I felt I was able to establish and maintain a consistent rhythm of dialogue with the AWACS-DTS system during the ACT-1 mission.	①	②	③	④	⑤	⑥	⑦	⑧
8. I felt that the fighter element was reacting appropriately to my communications.	①	②	③	④	⑤	⑥	⑦	⑧
9. I used the map display to effectively control and advise the fighters.	①	②	③	④	⑤	⑥	⑦	⑧
10. The appropriate level of feedback was provided by the facilitator before and after the mission.	①	②	③	④	⑤	⑥	⑦	⑧

6/11/2007

- | | | | | | | | | |
|--|---|---|---|---|---|---|---|---|
| 11. Overall, I was satisfied with the ease of completing my tasks during the mission. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 12. Overall, I was satisfied with the support information (online-line help, error messages, scenario / mission data) provided for the ACT-1 mission. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 13. Overall, I am satisfied with the amount of time it took to complete the ACT-1 training session. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 14. The fighter elements performed as I expected. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 15. Eglin Mission Control and the fighters never interrupted my communication. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 16. I believe the system wanted me to do or say something that was not appropriate to the mission. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 17. I had appropriate control over the fighters. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 18. Do you believe the behaviors or actions were supported or reinforced that would reduce the effectiveness of Air Weapons Officer. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 19. Proper skills needed to provide command and control to effectively support this mission are reinforced by the AWACS DTS. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 20. What did you <i>like</i> the most about the system mission? | | | | | | | | |
| 21. Do you believe the system would have been useful during your training? | | | | | | | | |
| 22. What kind of typical CBT provided level of support (online help, performance feedback) have you used? Which one of these would you like to have during this mission? | | | | | | | | |

6/11/2007

23. What did you *dislike* the most about the system mission?

24. Was there any particular point(s) in the TI-3 mission where you felt extremely frustrated – more so than at other points in the session? If so, briefly describe this instance(s) and the reason or cause for frustration.

25. Do you have any suggestions for improving the realism of the mission?

26. Did you experience confusion over which call sign or aircraft were simulating the friendly 2-ship?

27. Did requests for the SWAT score during the 2-ship separation disrupt your mission activities?

6/11/2007

**ETTAP Evaluation in Synthetic and Human Forces in a Research
Environment (SAFIRE)**

Evaluation of the AWACS Dialogue Training System (DTS)

Summary Evaluation Questionnaire

Instructions: Please indicate your level of agreement or disagreement with the statements below. 1 = Completely Disagree, 4 = No Opinion, 7 = Completely Agree, or Not Applicable (N/A). If you would like to provide greater detail on any of your answers, please do so by using the attached blank sheets. Be sure to indicate which questionnaire item for which you are providing additional information (e.g., Item A2, B6, etc.)

A. Ease of Learning

- | | | | | | | | | N/A |
|--|---|---|---|---|---|---|---|-----|
| 1. The voice recognition system was easy to learn. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 2. The AWACS Dialogue Training System (DTS) was easy to learn. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 3. The map display was easy to learn. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 4. The voice recognition/synthesis was easy to learn. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |

B. Ease of Use

- | | | | | | | | | |
|--|---|---|---|---|---|---|---|---|
| 1. The voice recognition systems was easy to use. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 2. The AWACS Dialogue Training System (DTS) was easy to use. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 3. The map display was easy to use. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 4. The voice recognition/synthesis was easy to use. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 5. I found it easy to get the system to do what I wanted it to do. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 6. I found the system to be flexible. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 7. The system allowed me to work at my own pace. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |

6/11/2007

C. Training Value

- | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1. I believe this system is capable of providing valuable training. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 2. This experience was positive. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 3. The system provided feedback to me on my progress throughout the mission. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 4. The feedback I received from the system was always appropriate to the situation and my performance. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 5. The feedback I received from the system was consistent throughout the mission. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 6. The system provided numerous points in the training for me to usefully interact with it. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 7. The AWACS-DTS training was as effective as traditional training received for a Weapons Director (WD) | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |

D. Naturalness of Voice Interaction

- | | | | | | | | | |
|--|---|---|---|---|---|---|---|---|
| 1. Voice interaction with the system seemed very "natural". | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 2. "Natural" voice interaction is important to promote effective training. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 3. My conversation with the system was realistic. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |
| 4. My radio transmissions were stepped-on only a few times. | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ | ⑧ |

E. General Questions

1. Please provide your assessment of the overall training value of the system to support AWACS Weapons Director/AWO training?

6/11/2007

2. When might a system like the AWACS DTS be useful?
3. Did you believe the system wanted you to say something that was not appropriate. If so, when?
4. I was able to evaluate the quality and realism of the simulated mission environment both for the communication and behaviors?
5. From your experience with today's technologies (voice recognition, map displays) today, could you see *additional* future uses for this system in the Air Force (i.e., non-AWACS related training)? If so, for what type of training?
6. What difference would you expect in the training experience obtained in this system vs. actual training experience?
7. Using the table below, please note the features of the current system that, in your opinion, are *most* in need of improvement in order to provide realistic training mission

6/11/2007

environment. For each feature listed, please indicate whether the amount of improvement required is "Extensive", "Moderate", "Slight", or "None".

Feature	Extensive	Moderate	Slight	None
1. Voice Recognition				
2. Audio sound quality				
3. Control Stations (Map Viewer)				
4. Fighter Action				
5. Eglin Mission Control system				
6. Voice recognition training				
7. Voice recognition usage				
8. Other (Specify)				
9. Other (Specify)				
10. Other (Specify)				

8. For each item checked as "Moderate" or "Extensive" in the table above, please describe the nature of the required improvement (in your opinion).

APPENDIX E. CONSENT FORM

INFORMATION PROTECTED BY THE PRIVACY ACT OF 1974

Informed Consent Document
For
AWACS-Dialogue Training System Evaluation
AFRL/HECP, WPAFB, OH, Building 33

Principal Investigator: Michael W. Haas, Principal Electronics Engineer, DSN 785-8768, AFRL/HECP, michael.haas@wpafb.af.mil

Associate Investigators: Mr. Robert Wnek, Human Factors Engineer, (937) 431-5898, Robert.wnek@ngc.com
Ms. Michael Salyer, Human Factors Engineer, (937) 431-5898, Michael.salyer@ngc.com

1. **Nature and purpose:** You have been offered the opportunity to participate in the "AWACS-Dialogue Training System(AWACS-DTS) Evaluation" research study. Your participation will occur in room B-07, Building 248, Area B, Wright-Patterson AFB. This facility contains a set of distributed simulations digitally linked together using distributed simulation technologies.

The purpose of this research is to determine if the AWACS-DTS has potential for improving the training of future AWACS weapon directors.

The data collection time requirement for each volunteer participant is anticipated to be a total of 4 hours. Training for each subject will be tailored to individual needs and progress on a self-paced schedule. A total of approximately 12 subjects will be enrolled in this evaluation. You will be required to report to the experimenter that you possess normal or corrected to normal visual acuity and normal hearing to be eligible for participation in this study. You will also need to possess a current SECRET clearance and be experienced with military air controlling operations.

2. **Experimental procedures:** If you decide to participate, you will be asked to provide information such that your security clearance can be verified by the AFRL/HEX security office. You will sit at a desktop workstation and perform simulated control over air intercepts. The simulation will be run at the Secret level. You will be asked to complete surveys to provide your feedback on your experience using the AWACS-DTS. Testing will last 4 hours and will be done in normal lighting conditions in a secured vault. You will be seated during the data collection trials. You will be offered a rest period midway through the testing period but can request a break at any time during the experiment. You may withdraw this consent at any time and discontinue further participation in this evaluation without prejudice to your entitlements. Also understand that the medical monitor of this evaluation

may terminate your participation in this evaluation if she or he feels this to be in your best interest.

3. **Discomfort and risks:** There is minimal risk and/or discomfort associated with performing this task. Mild postural fatigue and eye strain has been shown to be prevalent in continued computer usage but is normally alleviated with rest breaks.
4. **Precautions for female subjects:** There are no special precautions for female subjects.
5. **Benefits:** For active duty participants, there are no additional direct benefits to you for participation.
6. **Compensation:** No additional compensation will be provided for your participation.
7. **Alternatives:** Choosing not to participate is an alternative to volunteering for this evaluation.
8. **Entitlements and confidentiality:**
 - a. Records of your participation in this evaluation may only be disclosed according to federal law, including the Federal Privacy Act, 5 U.S.C. 552a, and its implementing regulations. Your personal information will be stored in a locked cabinet in an office that is locked when not occupied. Electronic files containing your personal information will be password protected and stored only on a DoD server. It is intended that the only people having access to your information will be the researchers named above and the AFRL Wright Site IRB, the Air Force Surgeon General's Research Compliance office, the Director of Defense Research and Engineering office or any other IRB involved in the review and approval of this protocol. When no longer needed for research purposes your information will be destroyed in a secure manner (shredding). Complete confidentiality for military personnel cannot be promised because information bearing on your health may be required to be reported to appropriate medical or command authorities.
 - b. You understand your entitlements to medical and dental care and/or compensation in the event of injury are governed by federal laws and regulations, and that if you desire further information you may contact the base legal office (88 ABW/JA, 257-6142 for Wright-Patterson AFB). You may contact the medical monitor, Dr. Jeff Bidinger, Maj., USAF, of this research evaluation at (937) 656-5449.
 - c. If an unanticipated event (medical misadventure) occurs during your participation in this evaluation, you will be informed. If you are not competent at the time to understand the nature of the event, such information will be brought to the attention of your next of kin.

Next of Kin or Health Care Agent if needed,

Name _____, Phone# _____.

- d. The decision to participate in this research is completely voluntary on your part. No one may coerce or intimidate you into participating in this program. You are participating because you want to. Michael Haas, or an associate, has adequately answered any and all questions you have about this evaluation, your participation, and the procedures involved. Michael Haas can be reached at (937) 255-8768. Michael Haas, or an associate will be available to answer any questions concerning procedures throughout this evaluation. If significant new findings develop during the course of this research, which may relate to your decision to continue participation, you will be informed. You may withdraw this consent at any time and discontinue further participation in this evaluation without prejudice to your entitlements. The medical monitor of this evaluation may terminate your participation in this evaluation if she or he feels this to be in your best interest. If you have any questions or concerns about your participation in this evaluation or your rights as a research subject, please contact Major Jeff Bidinger at (937) 656-5449 or jeffrey.bidinger@wpafb.af.mil.
- e. You understand that your participation in this evaluation may be photographed, filmed or audio/videotaped. The audio/video data will be used for data analysis, data retrieval and backup purposes only. All audio/video media will be stored in a secure cabinet for up to 5 years in room B-07, Building 248, WPAFB, OH. You understand that any release of records of your participation in this evaluation may only be disclosed according to federal law, including the Federal Privacy Act, 55 U.S.C. 552a, and its implementing regulations. This means personal information will not be released to unauthorized sources without your permission.
- f. YOU FULLY UNDERSTAND THAT YOU ARE MAKING A DECISION WHETHER OR NOT TO PARTICIPATE. YOUR SIGNATURE INDICATES THAT YOU HAVE DECIDED TO PARTICIPATE HAVING READ THE INFORMATION PROVIDED ABOVE.

Volunteer Signature _____ Date _____

Volunteer Name (printed) _____

Volunteer Social Security No. (Optional) _____

Advising Investigator Signature _____ Date _____

Investigator Name (printed) _____

Witness Signature _____ Date _____

Witness Name (printed) _____

We may wish to present some of the video/audio recordings from this study at scientific conventions or use photographs in journal publications. If you consent to the use of your image for publication or presentation in a scientific or academic setting, please sign below.

APPENDIX F. SUBJECTIVE WORKLOAD ANALYSIS TECHNIQUE

Mental workload refers to how hard you work to accomplish some tasks, or an entire job. The workload imposed on you at any point in time consists of various dimensions which contribute to the subjective feeling of workload. For this evaluation, we used the Subjective Workload Assessment Technique (SWAT) to assess workload

SWAT cards were sorted in order of the perception of increasing workload on a person. During the missions the subjects provided a SWAT score based on the opinion of the mental workload required to perform the task. The score consisted of the three dimensions: (1) Time Load, (2) Mental Effort Load, and (3) Psychological Stress Load.

APPENDIX G. SUMMARY OF PARTICIPANT RATINGS

Summary of Participant Ratings - TI-3 Mission

No.	Question	Participant/Subject							Avg.	SDev
		1	2	3	4	5	6	7		
1	I completed the TI-3 training session and aircraft did not go autonomous.	2	1	1	1	1	5	1	1.71	1.60
2	The TI-3 training with the AWACS-DTS was realistic.	5	5	1	5	4	5	1	3.71	1.60
3	The TI-3 mission was effective.	2	5	1	3	4	4	2	3.00	1.47
4	Verbal communications / responses from the AWACS-DTS during the TI-3 mission were consistently clear?	5	1	6	6	1	2	1	3.14	2.43
5	Verbal communications / responses from the AWACS-DTS during the TI-3 mission were consistently appropriate given the situation and time in which they occurred?	1	5	6	6	4	2	1	3.57	2.10
6	Verbal communications / responses from the AWACS-DTS during the TI-3 mission were consistently timely?	1	5	6	3	6	1	2	3.43	2.34
7	I felt I was able to establish and maintain a consistent rhythm of dialogue with the AWACS-DTS system during the TI-3 mission.	1	6	1	2	2	3	1	2.29	1.87
8	I felt that the fighter element was reacting appropriately to my communications.	1	2	1	2	1	1	2	1.43	0.52
9	I used the map display to effectively control and advise the fighters.	6	6	6	3	7	4	5	5.29	1.51
10	The appropriate level of feedback was provided by the facilitator before and after the mission.	5	5	7	7	7	7	7	6.43	1.03
11	Overall, I was satisfied with the ease of completing my tasks during the mission.	3	5	1	6	4	2	2	3.29	1.87
12	Overall, I was satisfied with the support information (online-line help, error messages, scenario / mission data) provided for the TI-3 mission.	3	3	6	4	6	2	2	3.71	1.67
13	Overall, I am satisfied with the amount of time it took to complete the TI-3 training session.	5	5	6	7	6	2	5	5.14	1.72
14	The fighter elements performed as I expected.	2	3	6	2	3	1	2	2.71	1.72
15	Eglin Mission Control and the fighters never interrupted my communication.	1	3	1	3	2	1	1	1.71	0.98
16	I believe the system wanted me to do or say something that was not appropriate to the mission.	2	3	2	4	2	3	4	4.14	0.82
17	I had appropriate control over the fighters.	1	3	1	2	1	1	2	1.57	0.84
18	Do you believe the behaviors or actions were supported or reinforced that would reduce the effectiveness of Air Weapons Officer.	4	3	7	5	3	1	2	3.43	2.04
19	Proper skills needed to provide command and control to effectively support this mission are reinforced by the AWACS DTS.	3	3	1	3	6	1	1	2.57	1.83

**Summary of Participant Ratings -
ACT-1 Mission**

No.	Question	Participant/Subject							Avg.	SDev
		1	2	3	4	5	6	7		
1	I completed the ACT-1 training session and aircraft did not go autonomous.	1	1	1	3	1	4	1	1.71	1.33
2	The ACT-1 mission was realistic.	6	7	2	3	6	3	4	4.43	2.07
3	The ACT-1 mission was effective.	3	7	1	1	6	3	4	3.57	2.51
4	Verbal communications / responses from the AWACS-DTS during the ACT-1 mission were consistently clear?	5	5	6	6	1	1	2	3.71	2.37
5	Verbal communications / responses from the AWACS-DTS during the ACT-1 mission were consistently appropriate given the situation and time in which they occurred?	2	5	6	6	3	4	2	4.00	1.63
6	Verbal communications / responses from the AWACS-DTS during the ACT-1 mission were consistently timely?	1	6	6	2	1	2	2	2.86	2.37
7	I felt I was able to establish and maintain a consistent rhythm of dialogue with the AWACS-DTS system during the ACT-1 mission.	1	6	1	1	2	2	2	2.14	1.94
8	I felt that the fighter element was reacting appropriately to my communications.	2	6	1	1	3	1	3	2.43	1.97
9	I used the map display to effectively control and advise the fighters.	6	7	6	2	7	7	7	6.00	1.94
10	The appropriate level of feedback was provided by the facilitator before and after the mission.	5	3	7	7	7	7	7	6.14	1.67
11	Overall, I was satisfied with the ease of completing my tasks during the mission.	4	5	1	5	2	4	4	3.57	1.64
12	Overall, I was satisfied with the support information (online-line help, error messages, scenario / mission data) provided for the ACT-1 mission.	3	2	6	4	4	4	4	3.86	1.33
13	Overall, I am satisfied with the amount of time it took to complete the ACT-1 training session.	5	7	6	7	6	4	5	5.71	1.17
14	The fighter elements performed as I expected.	2	7	6	2	5	1	3	3.71	2.48
15	Eglin Mission Control and the fighters never interrupted my communication.	2	2	1	1	3	1	2	1.71	0.82
16	I believe the system wanted me to do or say something that was not appropriate to the mission.	4	1	1	6	3	4	4	3.71	1.94
17	I had appropriate control over the fighters.	3	5	1	2	6	1	1	2.71	2.10
18	Do you believe the behaviors or actions were supported or reinforced that would reduce the effectiveness of Air Weapons Officer.	3	2	7	6	3	1	5	3.14	2.34
19	Proper skills needed to provide command and control to effectively support this mission are reinforced by the AWACS DTS.	1	6	1	2	5	1	2	2.57	2.25

**Summary of Participant Ratings -
Summary Evaluation**

No.	Question	Participant/Subject							Avg.	SDev
		1	2	3	4	5	6	7		
A1	The voice recognition system was easy to learn.	6	6	6	6	6	7	7	6.29	0.41
A2	The AWACS Dialogue Training System (DTS) was easy to learn.	5	6	6	6	6	6	4	5.57	0.41
A3	The map display was easy to learn.	6	6	6	6	7	5	7	6.14	0.63
A4	The voice recognition/synthesis was easy to learn.	5	6	6	5	6	6	7	5.86	0.52
B1	The voice recognition systems was easy to use	3	5	6	5	6	2	7	4.86	1.64
B2	The AWACS Dialogue Training System (DTS) was easy to learn.	3	6	6	1	6	3	4	4.14	2.14
B3	The map display was easy to learn.	5	6	6	6	7	7	6	6.14	0.75
B4	The voice recognition/synthesis was easy to learn.	4	6	6	4	6	2	6	4.86	1.63
B5	I found it easy to get the system to do what I wanted it to do.	2	5	1	1	6	1	4	2.86	2.12
B6	I found the system to be flexible.	2	2	1	1	3	1	4	2.00	1.15
B7	The system allowed me to work at my own pace.	4	3	1	3	3	1	2	2.43	1.13
C1	I believe this system is capable of providing valuable training.	2	7	1	5	3	1	2	3.00	2.24
C2	This experience was positive.	2	7	2	3	3	4	6	3.86	1.95
C3	The system provided feedback to me on my progress throughout the mission.	2	3	1	1	1	1	2	1.57	0.79
C4	The feedback I received from the system was always appropriate to the situation and my performance.	0	5	1	1	1	5	1	2.00	2.08
C5	The feedback I received from the system was consistent throughout the mission.	0	3	1	5	1	4	2	2.29	1.80
C6	The system provided numerous points in the training for me to usefully interact with it.	5	5	1	2	3	1	4	3.00	1.73
C7	The AWACS-DTS training was as effective as traditional training received for a Weapons Director (WD)	3	3	1	1	4	1	1	2.00	1.29
D1	Voice interaction with the system seemed very "natural".	3	5	4	1	5	1	5	3.43	1.81
D2	"Natural" voice interaction is important to promote effective training.	7	7	7	7	7	7	6	6.86	0.38
D3	My conversation with the system was realistic.	2	6	2	7	6	2	7	4.57	2.44
D4	My radio transmissions were stepped-on only a few times.	2	3	1	1	5	1	3	2.29	1.50
F1	Voice Recognition	3	1	0	3	2	2	3	2.00	1.15
F2	Audio sound quality	1	1	0	1	1	1	1	0.86	0.38
F3	Control Stations (Map Viewer)	0	1	0	1	1	1	2	0.86	0.69
F4	Fighter Action	1	1	0	2	3	3	3	1.86	1.21
F5	Eglin Mission Control system	2	3	0	3	3	3	2	2.29	1.11
F6	Voice recognition training	2	2	0	2	2	0	2	1.43	0.98
F7	Voice recognition usage	3	2	0	2	2	1	1	1.57	0.98