

Measuring Situation Awareness for Dismounted Infantry Squads

Donald R. Lampton
Army Research Institute
Orlando, FL
donald.lampton@us.army.mil

Mica R. Endsley
SA Technologies
Marietta, Georgia
mica@satechnologies.com

Michael T. Gately
ScenPro
Richardson, TX
mgately@scenpro.com

Joseph V. Cohn
Naval Research Laboratory
Washington, DC
cohn@itd.nrl.navy.mil

Jared Freeman
Aptima
Washington, DC
freeman@aptima.com

Glenn A. Martin
University of Central Florida
Orlando, FL
martin@ist.ucf.edu

ABSTRACT

Situation Awareness (SA) is a fundamental aspect of contemporary military operations. Our paper provides the background for and describes the first evaluation in a planned three-year program to develop new methods of measuring and providing performance feedback on team SA. Our initial focus is on small teams of dismounted infantry conducting training exercises using virtual simulations. A three-pronged approach to measuring SA is planned. The first employs systematically administered SA probes scored in real-time during the team exercise. This approach will be augmented by trainer ratings of SA performance. The second approach involves the capture and analysis of voice communications among team members. A new system incorporates automated analysis of voice communications, including communications flow (sender, receiver, and duration) and content (category of transcribed communication). The third approach involves monitoring the networked simulation data stream during an exercise. Predefined rules combine assessments of movement, orientation, and weapon use to produce measures of decision making and SA. The three approaches to measuring SA were assessed in the context of an After Action Review (AAR) for squad-level exercises. Squads conducted urban missions using the Squad Synthetic Environment at the Fort Benning Soldier Battle Lab. The exercises will employ scenarios developed specifically to support the measurement of the squad's ability to gain, maintain, and apply SA. In future years the SA measures will be used to provide performance feedback to the trainees during an AAR following each exercise, and will support training evaluation and management, as well as SA training research. This research involves collaboration of the Army Research Institute, Office of Naval Research, the University of Central Florida Institute for Simulation and Training, and Army/OSD SBIR contractors: SA Technologies, Aptima, and ScenPro.

ABOUT THE AUTHORS

Donald R. Lampton is a Research Psychologist with the Army Research Institute for the Behavioral and Social Sciences (ARI). For the last eighteen years he has been with the ARI Simulator Systems Research Unit in Orlando, FL. He is currently conducting research on training applications of Virtual Reality technology. He is co-developer of the Virtual Environment Performance Assessment Battery (VEPAB), the Fully Immersive Team Training (FITT) system, and the Dismounted Infantry Virtual After Action Review System (DIVAARS).

LCDR Joseph V. Cohn, Ph.D., is designated a Naval Aerospace Experimental Psychologist. Currently, Lieutenant Cohn serves as the Lead, Requirements and Training for Virtual Environments (VE) at the Naval Research Laboratory where he is developing a comprehensive theoretical framework for establishing Training Transfer. As such, he is directing the evaluation of deployable VE training systems for the Naval Aviation and Expeditionary Warfare communities and providing direct training and operational Fleet support to these communities.

Mica R. Endsley, Ph.D., is President of SA Technologies in Marietta, Georgia. She received a Ph.D. in Industrial and Systems Engineering from the University of Southern California. She has published extensively in the area of situation awareness and is the co-author of *“Designing for Situation Awareness”*.

Jared Freeman, Ph.D., is Vice President for Research at Aptima, Inc., a Human-Centered Engineering firm located in Woburn, Massachusetts and Washington, D.C. Dr. Freeman has served as P.I. for many assessment and training projects at Aptima, and he is the author of 60 articles and book chapters on these and related topics. Dr. Freeman received a Ph.D. in Human Learning and Cognition from Columbia University, and an M.A. in Educational Technology from Teachers College, Columbia University.

Michael T. Gately is the Chief Technical Officer of ScenPro and was the principal investigator for the ViSSA and SA-STAT programs. He holds a BS in Systems Engineering and MS in Computer Science from Georgia Institute of Technology. He has taught a variety of courses in industry and has been an Adjunct Associate Professor at UT-Austin and UT-Dallas.

Glenn A. Martin is a Senior Research Scientist at the University of Central Florida's Institute for Simulation and Training (IST) where he leads the Interactive Realities Laboratory, pursuing research in multi-modal, physically realistic, networked virtual environments and applications of Virtual Reality technology. He has led numerous projects related to virtual reality including a testbed for evaluating VR for training uses, an investigation of network architectures, a library for VR applications, and the Dismounted Infantry Virtual After Action Review System (DIVAARS).

Measuring Situation Awareness for Dismounted Infantry Squads

Donald R. Lampton
Army Research Institute
Orlando, FL
donald.lampton@us.army.mil

Mica R. Endsley
SA Technologies
Marietta, Georgia
mica@satechnologies.com

Michael T. Gately
ScenPro
Richardson, TX
mgately@scenpro.com

Joseph V. Cohn
Naval Research Laboratory
Washington, DC
cohn@itd.nrl.navy.mil

Jared Freeman
Aptima
Washington, DC
freeman@aptima.com

Glenn A. Martin
University of Central Florida
Orlando, FL
martin@ist.ucf.edu

OVERVIEW

Dismounted infantry operations in urban areas are difficult, frequent and numerous, critical to success, and bear heavy consequences for failure. Preparing units for urban missions is one of the greatest challenges currently facing the training and simulation community (Lampton, Clark, & Knerr, 2003). Dismounted Soldier virtual simulation capabilities can help meet this challenge by supporting training, mission rehearsal, and concept development.

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) Simulator Systems Research Unit (SSRU) has an ongoing program of behavioral research to support the development and use of dismounted Soldier virtual simulations. This program includes three ongoing Office of the Secretary of Defense (OSD)/Army Small Business Innovation Research (SBIR) Phase II contracts.

VE After Action Review (AAR) systems are very good at quickly and convincingly displaying “what” happened during a mission exercise. The more difficult issues of “why it happened” and “how to do better” frequently require discussion of Situation Awareness (SA). SA has been a valuable concept for performance measurement and training in other areas. The goal of this program is to produce a system for measuring SA at the squad level that is unobtrusive, in contrast to a “freeze exercise” procedure that might be effective with a single trainee but would be disruptive for a squad level exercise with several trainees, OPFOR role players, and computer generated forces. Two of the SBIR contracts directly address measurement of

Situation Awareness (SA) during squad-level exercises conducted in immersive Virtual Environments (VEs). The third is developing an automated system for measuring communication performance. VE systems can readily capture, analyze, and display individual movement and weapons use data. In marked contrast, voice communications can be recorded, but transcription and analysis is slow and tedious. Thus, what should be one of our richest sources of information about critical team processes and SA is seldom amenable to timely analysis.

All three contractors were tasked with developing performance measurement systems that would support both feedback to the trainees during an AAR conducted soon after the completion of a VE exercise, and more in-depth analyses to support research and training management.

Prototypes of all three systems were evaluated by infantry squads at the Soldier Battle Lab, Fort Benning, GA in June 2005. Soldier assessment enhanced the effectiveness and efficiency of the system development process.

This research is being conducted in collaboration with the Office of Naval Research (ONR) Virtual Technologies and Environments (VIRTE) program. The involvement of the VIRTE program will make sure that the results are immediately applicable to Marine training and simulation, and generalizable to other domains as well.

The remainder of our paper is organized as follows. First, an overview of the R&D programs is presented

to establish the context of the overall effort. Next, each of the three measurement systems is described in detail. The Soldier assessment of the systems at the Soldier Battle Lab is briefly described. Conclusions and plans for future research are presented.

OFFICE OF NAVAL RESEARCH VIRTUAL TECHNOLOGIES AND ENVIRONMENTS (VIRTE) PROGRAM

Virtual Environment training systems offer a unique solution for a range of training requirements. Yet while the technologies supporting the advancement of these systems continue to be refined, the degree to which these systems can support training and enhance real world performance is often overlooked or, worse, simply 'assumed'. The Virtual Technologies and Environments (VIRTE) program set out to address this challenge in three stages. The first stage, Virtual Environment Expeditionary Warfare (VE EW) endeavored to demonstrate the degree to which a Human Centric Simulator Design Model could be developed, validated and utilized to rapidly prototype Vehicle based VE training systems. The second stage, Virtual Environment Human Interface Technology (VE HIT), capitalizing on the success with which this model has had thus far, is currently focusing on non-Vehicle team based VE training systems, that is, systems that directly interface to individual users, rather than through the metaphor of a vehicular intermediary. The third stage, Multi-Operational Team Training Immersive Virtual Environments (MOTIVE) will synthesize the results of the earlier efforts in a unique, cross-platform multispectrum training environment that will include advanced training enhancement tools and strategies, and which will also lay the foundation for addressing training challenges of Joint Services, as well as Coalition forces.

Each of these stages, while endeavoring to demonstrate real world performance enhancement through exposure to VE, has its own unique evaluative metrics. For VE EW, these metrics included a heavy reliance on Usability Engineering principles to guide system development. For VE HIT, these metrics will shift to more advanced analyses of the minimum ecological validity that individual user interfaces must have, as well as to the development of a team performance evaluation metrics suite. For MOTIVE, metrics will

focus on both collective training, an advanced form of team training, as well as on an evaluation of the benefits of integrating performance enhancing tools within VE. The intent with this approach is to demonstrate a spiral development cycle not only for the production of actual systems but, also, in the collective understanding of what VE systems are (and are not) capable of accomplishing.

VIRTE emphasizes the merging of Training Science with real world applications. At each stage, basic research questions are identified and addressed within a system or a suite of systems that will transition to their respective DoD customers.

ARI VIRTUAL ENVIRONMENTS FOR SMALL UNIT LEADER TRAINING RESEARCH PROGRAM

Future Force small unit leaders and teams will operate in highly complex decision environments with technological tools and information capabilities never before available. Emerging Soldier system technologies, e.g. Future Force Warrior (FFW), promise to provide revolutionary advances in small unit effectiveness, but their potential will only be realized if leaders and Soldiers can be trained to take full advantage of the new capabilities and operational concepts. No guidelines, however, exist that specify what to train, how to train, and how to measure success of training for small unit leaders and teams to take better advantage of new Future Force capabilities, operational concepts, and tactics, techniques, and procedures.

ARI is conducting research to leverage advances in situation awareness training and measurement, computer gaming and simulation, and intelligent tutoring to meet these needs. We will determine the best methods for using wearable computers to train small unit leaders and teams. The final product will include new training methods and performance measures required to exploit new Future Force capabilities and high-tech equipment, and embedded training methods to accompany hardware and software.

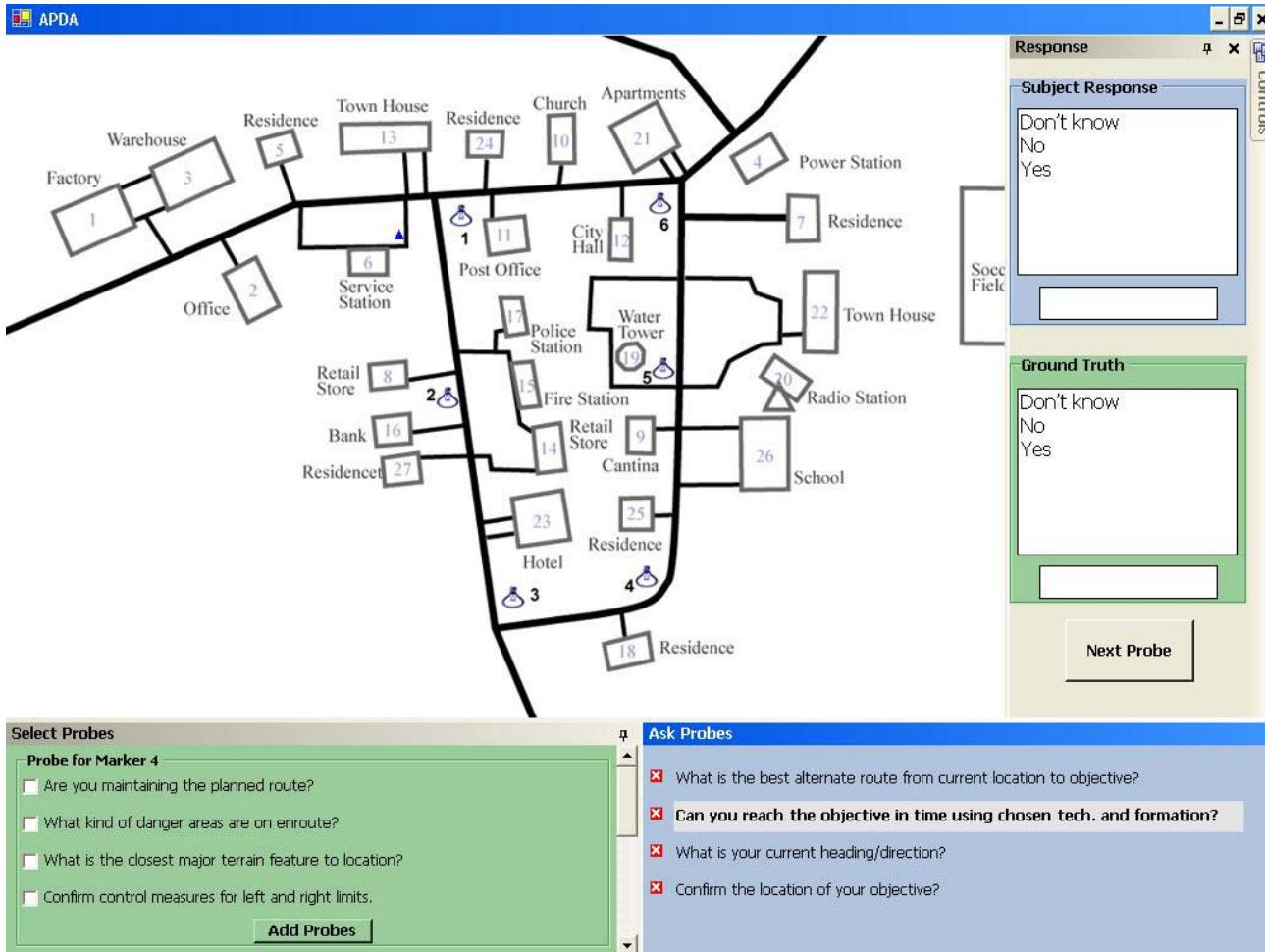


Figure 1. VESAARS Probe Selection and Administration Subsystem

VIRTUAL ENVIRONMENT SITUATION AWARENESS ANALYSIS AND REVIEW SYSTEM (VESAARS)

VESAARS is being developed by SA Technologies as a comprehensive system for measuring, and providing detailed feedback on, the SA of Soldiers participating in VE simulations. VESAARS is comprised of three complementary SA measurement approaches. Each of these measures was selected based on an extensive analysis of the validity, reliability and sensitivity of SA measures, and based on an assessment of their ability to provide a full range of information on SA for use in training. The administration and output of all three measures are controlled through a single integrated system.

Real-time SA probe measure. The real-time SA probe measure provides questions or probes to Soldiers during training that tap into their SA of ongoing situations. The real-time probing approach is based on

a cognitive task analysis of squad-level operations to identify Soldier SA requirements that form the probes to be presented.

A customized probe database has been developed for each Military Operations in Urban Terrain (MOUT) training scenario, and is associated with specific MOUT events (and squad goals) and locations as part of the training mission. The VESAARS Probe selection and administration subsystem is shown in Figure 1. Probes are verbally administered by radio to a squad leader by a confederate platoon leader using VESAARS, which monitors the exact position of a squad leader's avatar in the virtual training environment. All probes are classified according to the three levels of SA described by Endsley (1995). Examples of probes include: (1) Can you confirm enemy threat is in your sector? (Level 1- perception); (2) Can the enemy engage you from where you are? (Level 2-comprehension); and (3) How long can you remain in your current position without endangering

the squad? (Level 3-projection). The probes are written in the form expected as part of normal communications for a squad leader. VESAARS allows the platoon leader to record Soldier responses to the probes using a touch-screen for quick entry. The platoon leader also enters the “ground truth” of the training simulation. These data are used to automatically score a Soldier’s SA and to generate overall SA scores.

Electronic SA behavior rating measure (ESABM).

The second measure was developed based on prior research on a Situation Awareness Behaviorally Anchored Rating Scale (SABARS) (Strater, Endsley, Pleban & Matthews, 2001). SABARS was developed to subjectively evaluate SA of platoon leaders in military training exercises. The items of the original SABARS were modified to measure SA for squad leaders and the squad members. The new ESABM includes 27 critical Soldier SA behaviors in MOUT, such as:

- solicits information from fire-team leaders;
- communicates key information to platoon leader;
- performs recon to assess terrain and situation; and
- projects future possibilities and creates contingency plans.

All ratings are made on 5-point scales ranging from “very poor” to “very good” and are entered into VESAARS which automatically computes aggregate SA behavior scores in real-time, and records the exact time in a scenario at which a rating was made in order to later associate expert assessments with specific scenario events during AARs.

SA measure of team communication and coordination. The third measure of VESAARS provides a process measure of Soldier situation assessment allowing for insight into how SA develops over the course of training simulation trials. The approach to this measure was based on Wright and Kaber’s (2003; 2005) recent adaptation of team communication and coordination measures for assessing team situation assessment behaviors. Expert observers analyze Soldier verbalizations during virtual training exercises by categorizing and counting specific communications and identifying trainee attempts at situation assessment. The categories of communication defined as part of the measure are specific to squad-level MOUT and include:

- seeking updates on battlefield situation;
- attempting to identify problems or need for action;
- communication to prevent errors in performance;
- statements suggesting deviations in SA among team members;
- statements regarding unit status;

- integration of information from multiple sources; and
- attempting to prioritize tasks accurately.

The experts use VESAARS to electronically count the communications, rating the quality of situation assessments as “good” or “bad”. VESAARS also calculates totals for each situation assessment behavior, and specific elements of MOUT. This is used as a basis for making an overall rating of Soldier skill in developing and maintaining SA.

Feedback and Evaluation. An SA reporting module summarizes the results across the three VESAARS measures based on the three levels of SA and various elements of MOUT. The reporting module provides detailed feedback on each specific measure including a mission knowledge assessment based on the real-time probe technique, an assessment of the appropriateness of specific SA behaviors based on the ESABM results, and an assessment of the quality of specific situation assessment communications on the elements of MOUT. The reporting module provides the capability for an AAR leader to retrieve a dialog presenting: (1) the specific (Level 1, 2 and 3) SA probes that were posed to a squad leader, (2) the Soldier responses to the probes; and (3) the “ground truth” for probes, or “good” SA at the time of a probe. In this way, the system provides for explanatory power along with the high-level SA score summaries for each measure as part of the system.

THE VIRTUAL SOLDIER SKILLS ASSESSOR (ViSSA) AND INTEGRATED SITUATIONAL AWARENESS TOOLKIT (SA-STAT)

Developed by ScenPro, Inc., ViSSA and SA-STAT comprise an integrated software suite that aids in the assessment of dismounted infantry VE exercises. ViSSA is designed to assess small unit team leader decision-making skills and was developed under previous SBIR funding. SA-STAT adds unobtrusive individual and team-level SA assessment.

ViSSA operates by eavesdropping on the Protocol Data Unit (PDU) packets (entity status information) sent between virtual entities. It creates an Object Model of the virtual world and then evaluates entity actions against a set of rules. In particular, rules consist of Event/Condition/Action (E/C/A) triplets. The Events are either the arrival of certain PDUs or the expiration of a timer. The Conditions are an arbitrary set of IF clauses using data from the Event and the

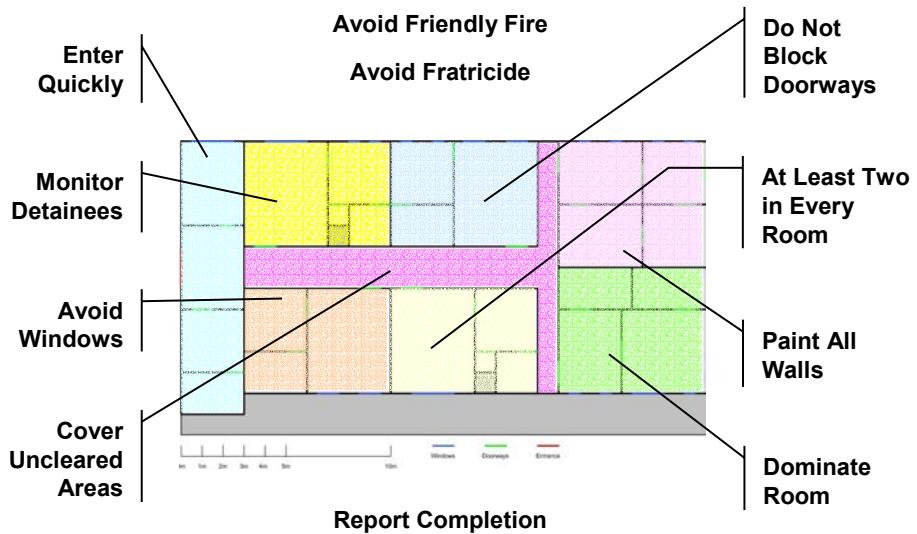


Figure 2. Floor Plan for ScenPro Building Clearing Exercise and Examples of Performance Measures

Object Model. If all of the clauses are true, one or more Actions are executed. Actions include setting variables, starting or canceling a timer, or marking the exercise timeline with a good or poor decision marker. At a low level, E/C/A rule authoring is similar to writing Excel macros. ViSSA contains rule templates that aid in the construction of E/C/A rules.

As entities move within the virtual world, the Object Model is updated with a wide variety of information including location, velocity, orientation, posture, health, rounds fired, and radio status. These Object Model values, plus timers, knowledge of lines and areas overlaid onto the terrain, entity role and affiliation, and user-defined variables are combined in a powerful, comprehensive assessment system.

SA-STAT is a flexible tool for unobtrusive, real-time SA assessment (based on information available and limitations of the VE) that can support any formal SA measurement methodology. SA-STAT operates using Situational Awareness Measurements (SAM). A SAM consists of a Starting Trigger, a Measurement, and an Ending Trigger. Starting and Ending Triggers are implemented as Events and Conditions and used to constrain the spatio-temporal applicability of the SAM. For example, a SAM may be restricted to the period of time just after the enemy begins firing.

ViSSA uses the SAMs to automatically and objectively assess an entity's actions. Figure 2 shows the floor plan for a VE building model for use in assessing squad

building-clearing performance. Examples of performance measures that ViSSA SA-STAT can be programmed to score automatically based on the building model are also shown in the figure.

SAMs can also be used to initiate one of three types of questionnaires (subjective measures): naturalistic Soldier probes, observer/controller questionnaires, and Soldier AAR questionnaires (completed after the exercise). Naturalistic Soldier probes require the observer/controller, as a role player, to use the radio system to prompt the Soldier for information and answer questions concerning the Soldier's response.

SAMs are assigned to an individual or a team. As SAMs are triggered, a score is generated for the individual or team. The observer/controller can see the scores in real time. The observer/controller has the ability to supplement or override computer assessments.

ViSSA/SA-STAT operates under Windows and is comprised of three software modules: Logger, Audio Logger, and Assessment. The entire ViSSA/SA-STAT system is operated by a single observer/controller sitting at the Logger console. The two loggers synchronously record an exercise's PDU traffic. This PDU traffic is replayed during the AAR to allow the Soldiers being assessed to see and hear their movement and behavior on 2D and 3D displays. The Logger calculates and displays 60 different entity-level statistics during the exercise. After an exercise these

statistics can be exported to Microsoft Excel and can be used to create reports in Microsoft Word. The Audio Logger works with ASTi's audio simulator and outputs to the computer's speakers. The Event/Condition/Action rules are implemented using Microsoft COM Automation, Microsoft Scripting Engine, the VB Script language, and the ViSSA/SA-STAT Object Model.

ViSSA/SA-STAT has the ability to host an AAR immediately after an exercise. The Logger user interface allows the AAR Leader to quickly focus attention on significant events in the exercise where the team showed good or poor decision making or situational awareness.

AUTOMATED COMMUNICATIONS ANALYSIS SYSTEM (AUTOCAS)

Some of the best data for assessing SA, decision making, and leadership in team settings are also the least used. Verbal communications are ubiquitous and easy to capture. However, it has traditionally taken days or weeks to analyze these data and assess performance using them. As a result, communications analysis has become largely a research activity. Such research has produced valuable insights into the structure of expert knowledge concerning military operations (c.f., Freeman, Thompson, & Cohen, 2000), and patterns of discourse in successful teams (Entin & Serfaty, 1999). A more trainer-friendly and rapid form of communications analysis might greatly improve assessment and debriefing of SA and teamwork in training exercises.

Aptima developed the Automated Communications Analysis System (AutoCAS) to help trainers analyze and debrief team communications in MOUT exercises. AutoCAS provides a workbench on which trainers can build and present AARs concerning communications, and it performs rapid analyses of communications patterns to enrich these AARs. An additional subsystem classifies transcribed voice communications in categories of value to MOUT trainers and researchers.

As depicted in the interface (Figure 3), trainers can use AutoCAS to inspect (a) the time course of communications, (b) a transcript of communications, (c) the pattern of communications within and between trainees, and (d) descriptive statistics concerning these

patterns. Communications replay and window presence are managed with a control panel (not shown) and (e) a utility for bookmarking communications and related graphics. All windows can float (as shown) or dock within a master window, and expand or shrink to suit a user's preferences and screen space. Underlying these capabilities are two major forms of automated support to trainers: speech-to-text transcription and statistical analysis of communications patterns and content.

AutoCAS transcribes recorded speech from any of several standard audio file formats into text. The current version of AutoCAS integrates the Audio Mining Development System, manufactured by ScanSoft® Inc. to support this function. This component reads a digitized audio file, applies an acoustic model to perform speech recognition, and generates a time-stamped index of words and utterances in Extensible Markup Language (XML) format. AutoCAS then loads the XML file into its data repository and rapidly generates transcripts. Recognition is speaker-independent. Under good conditions the system achieves 90% accuracy without user training. Recognition accuracy can be refined prior to training, if necessary, by providing the system with textual communication samples from the domain, with which it estimates word and phrase probabilities.

Automated classification analysis assigns each communication to classes important in MOUT training and operations. Three sets of classes are used. The first set is the doctrinal METT-TC classification system for analyzing military situations. The acronym denotes mission, enemy, terrain and weather, friendly troops and support, time available, and civil considerations. The second classification scheme concerns general functions such as see/hear, plan, move, and others. The third scheme (Bowers, Jentsch, Salas, and Braun, 1998) defines communications-specific functions, such as command, question, and acknowledge.

To classify the transcribed communications, AutoCAS uses Bayesian Binary Regression, a supervised machine learning algorithm. This algorithm learns a set of classes from a manually coded library of communications, computes the probability that a new (training) communication belongs to each class of communication, and returns the name of the class to which the communication most likely belongs. This functionality operates off-line, and is not yet integrated into the AAR interface.

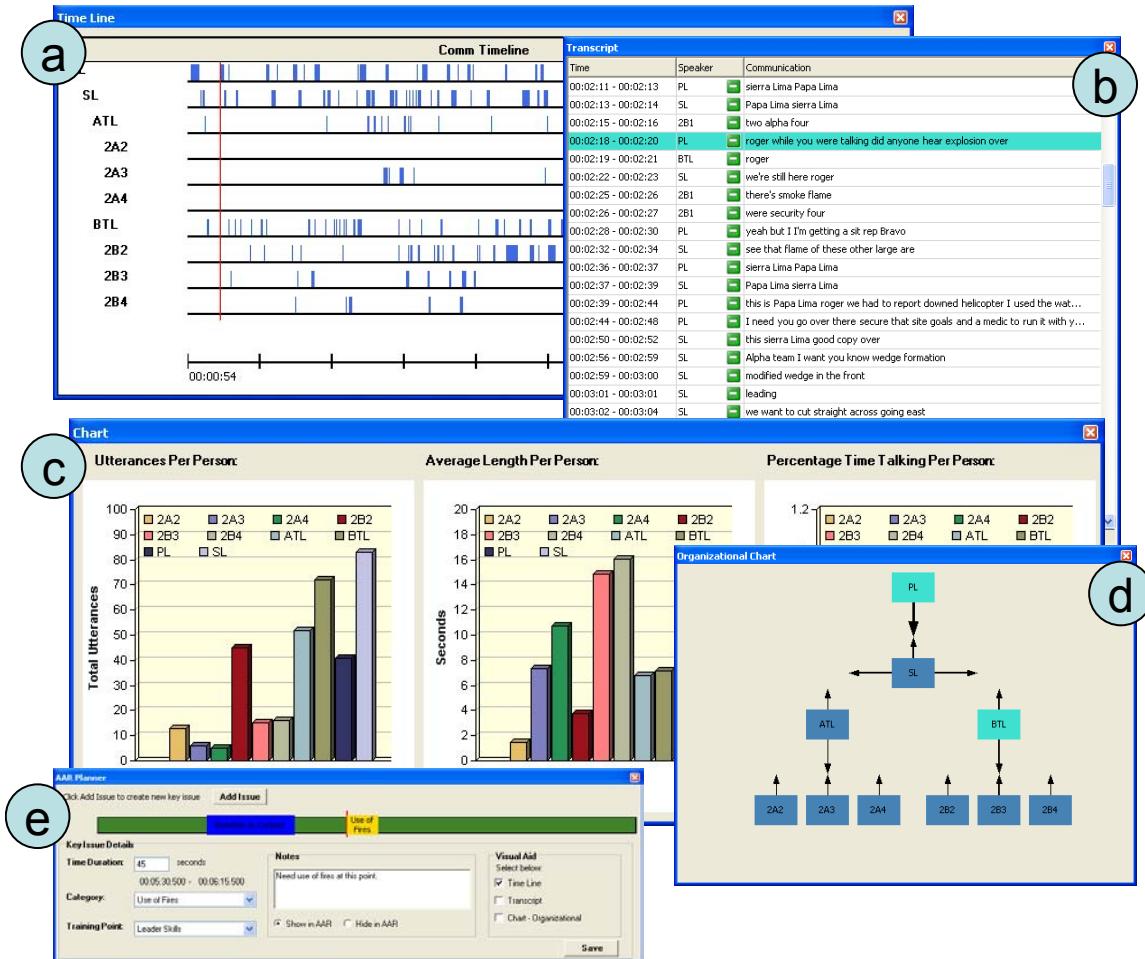


Figure 3. AutoCAS Communications Analysis Interfaces

In addition, AutoCAS helps trainers to generate and present AARs by enabling them to bookmark communications events and relevant graphs before the AAR, and to serially play those communications and display the graphs during the AAR.

The AutoCAS architecture is highly modular. It provides programming interfaces into which programmers can integrate different speech recognition and categorization tools to provide the functions above.

In sum, AutoCAS is designed to bring communication analysis out of the cloister of research and into training exercises. It does so by providing trainers with pattern analyses, transcriptions, and classification analyses of verbal communications data, as well as interfaces with which to author and deliver After Action Reviews of communications. These capabilities should improve the assessment of situation awareness and teamwork during training by tapping the rich stream of data –

verbal communications – through which shared situation awareness is achieved.

SYSTEM ASSESSMENT AT THE SOLDIER BATTLE LAB

VESSARS, ViSSA SA-STAT, and AutoCAS were evaluated with Soldiers using the Squad Synthetic Environment (SSE) at the Soldier Battle Lab to determine the extent to which they could capture their intended performance measures, and to identify ways to improve the systems. The SSE employs rear-screen projection systems to individually immerse each of a nine member squad into a VE model. A Distributed Interactive Simulation (DIS) network enabled the squad members to act as a team. The SSE represents many of the weapons and resources (smoke grenades and signal flares) that would be available to a squad during actual MOUT operations.



Figure 4. After Action Review with the Dismounted Infantry Virtual After Action Review System

Computer generated Semi-Automated Forces populated the VE exercises with friendly forces, heavy civilian pedestrian and vehicular traffic, and Opposing Forces.

Three squads participated in the assessment. Each squad followed the same two day schedule:

Day 1: Highly structured, systematic train-up on the simulators. Building clearing exercises. Three complex patrol missions and/or search missions.

Day 2: Building clearing exercise. Three complex patrol missions and/or search missions. Questionnaires and exit interviews.

Communications performance data were captured during all of the exercises. After every exercise, an AAR was conducted (Figure 4) using the Dismounted Infantry Virtual After Action Review System (DIVAARS) developed by SSRU and the University of Central Florida's Institute for Simulation and Training (IST) (Knerr et al., 2002). Simulator sickness

questionnaires were administered at the beginning, mid point, and conclusion of each daily sequence of exercises.

PRELIMINARY FINDINGS

The building clearing exercises effectively presented a comprehensive array of challenges selected from infantry and Special Weapon and Tactics (SWAT) training requirements. The ViSSA SA-STAT system succeeded in automatically scoring individual, fire team, and squad performance measures. Information gained from assessing the system with Soldiers led to the development of additional rules.

The VESAARS subsystem for selecting, administering, and scoring real-time probes worked very well. The probes integrated well with normal radio communications, and did not disrupt the temporal flow of the exercises. The subsystems for expert ratings of SA also worked well, but presented a challenging workload for the raters. The VESAARS subsystems

were tailored to the unique advantages and limitations of conducting exercises in VEs. However, they could be readily adapted for live training.

The MOUT scenarios, situated in the Shughart-Gordon VE, were well suited for testing the squads' abilities to gain and maintain Situation Awareness. Variations in OPFOR aggressiveness, and manipulations of SSE environmental effects (night versus day, clear versus rain) were effective in influencing exercise difficulty level.

The AutoCAS system was successful in capturing, and providing striking visual displays of, many aspects of voice communications performance across the course of an exercise. It was also effective for highlighting communications associated with specific critical events, for example when the squad was confronted by succession of command challenges. The results with the AutoCAS voice recognition component were less successful. Several approaches for improvement will be pursued.

As in previous evaluations at Fort Benning and Fort Campbell, the Soldiers judged DIVAARS to be a very effective AAR tool. What events happened during a mission could be determined quickly and decisively so that the AAR could focus on the difficult issues of why events occurred and how to do better the next time. Also as in previous evaluations, simulator sickness was not found to be a problem with the rear-screen projection systems.

These findings will be used to guide the next stages in the design, implementation, and fielding of the respective systems. Next year's activities will concentrate on refining the SA and communications measures to have the appropriate content and format for supporting AARs. A major challenge is the limited amount of time available to generate performance feedback before the AAR.

The following year will focus on the effectiveness of an AAR system integrating all the performance measures. The AAR procedure is a bedrock of Army and military training. However quantifying the effectiveness of AARs remains a challenge.

REFERENCES

Bowers, C.A., Jentsch, F., Salas, E., & Braun, C.C. (1998). Analyzing communication sequences for team training needs assessment. *Human Factors*, 40, 672-679.

Endsley, M. R. (1995). Toward a theory of situation awareness. *Human Factors*, 37(1), 32-64.

Entin, E.E & Serfaty, D. (1999). Adaptive team coordination. *Human Factors*, 41, 312-325.

Freeman, J.T., Thompson, B.T., & Cohen, M.S. (2000). Modeling and assessing domain knowledge using latent semantic indexing. *Special Issue of Interactive Learning Environments*, 8 (3), 187-209.

Knerr, B.W., Lampton, D.R., Martin, G.A., Washburn, D.A., & Cope, D. (2002). Developing an After Action Review System for Virtual Dismounted Infantry Simulations. *Proceedings of the 2002 Interservice/Industry Training, Simulation and Education Conference*. Orlando, FL.

Lampton, D. R., Clark, B. R., & Knerr, B. W. (2003). Urban Combat: the Ultimate Extreme Environment. *Proceedings of the first annual meeting of the Society for Human Performance in Extreme Environments*. Denver, CO.

Strater, L. D., Endsley, M. R., Pleban, R. J., & Matthews, M. D. (2001). *Measures of platoon leader situation awareness in virtual decision-making exercises* (Research Report No. 1770). Alexandria, VA: U.S. Army Research Institute for Behavioral and Social Sciences.

Wright, M.C. & Kaber, D. B. (2003). Team coordination and strategies under automation. In *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 553-557). Santa Monica, CA: Human Factors and Ergonomics Society.

Wright, M. C. & Kaber, D. B. (2005). Effects of automation of information processing functions on teamwork. *Human Factors*, 47(1), 50-66.