Developing Situation Awareness Metrics in a Synthetic Battlespace Environment

Jacqueline M. Curiel, Michael D. Anhalt
Alion Science and Technology
Fairfax, Virginia
{jcuriel,manhalt}@alionscience.com

John J. Tran, Ke-Thia Yao
Information Sciences Institute/USC
Marina del Rey, California
{jtran,kyao}@ISI.EDU

ABSTRACT

The Joint Forces Command (JFCOM) conducts Joint Urban Operation (JUO) exercises in synthetic battlespace using human-directed computer simulation tools such as Joint Semi-Automated Forces (JSAF) to support ongoing joint war-fighting efforts. A component of these experiments is that of human-in-the-loop (HITL) interactions where human players impact the outcome of the exercise. This is in contrast to Monte Carlo constructive experiments that only involve computer behavior. The need to objectively measure the effectiveness of human players and their interaction with the simulation environment requires quantitative metrics to supplement more qualitative observer-based judgments. Situation awareness (SA), a cognitive behavior captured in HITL experiments, involves the perception and comprehension of forces and events in a situation, and a prediction of their future status, Endsley (1995). Objectively measuring SA is drawing intense interest because this knowledge is crucial to successful decision-making processes (C2).

Building upon work presented at I/ITSEC 2004 (An Interdisciplinary Approach to the Study of Battlefield Simulation Systems, paper 1886), we adopt a cognitive-computational approach for measuring SA based on Situation Model theory. Situation models are complex mental representation of events. As events unfold, these mental representations must be updated to maintain an accurate representation. Prior research has demonstrated that situation models are updated along a number of dimensions. These dimensions reflect information about entities, space and time coordinates, participants' goals, and the causal relationships of events. We utilize the information encapsulated in SA objects (SAOs), recorded during the JUO exercises, to develop a tool that automatically monitors players' SA and evaluate the importance of these dimensions on situation awareness over the time course of the experiment and on the three levels of SA. Our findings have practical implications for subsequent training, product development, and extend the knowledge base of cognitive behavior.
.

ABOUT THE AUTHORS

Jacqueline M. Curiel is a research psychologist at Alion Science and Technology. She is also a co-founder of Behavioral Cognition and is a consultant to IdeaDaVinci, a technology incubator. Her prior academic experience includes teaching and research positions at the University of Texas at San Antonio and the University of Notre Dame, where she did her graduate work. Her published work has primarily included work on mental maps and situation models, the focus of her Master's thesis and doctoral dissertation.

John J. Tran is a researcher at the Information Sciences Institute, University of Southern California. He received both his BS and MS Degrees in Computer Science and Engineering from the University of Notre Dame, where he focused on Object-oriented software engineering, large-scale software system design and implementation, and high performance parallel and scientific computing. He has worked at the Stanford Linear Accelerator Center, Safetopia, and Intel. His current research centers on Linux cluster engineering, effective control of parallel programs, and communications fabrics for large-scale computation. Capt Tran is also a member of the 129th Rescue Wing at Moffett FAF, California.

Michael D. Anhalt is retired Navy Surface Line Commander with over 23 years of operational experience, including specialties in Amphibious Warfare, Surface, Undersea, and Strike Warfare, and tactical training. Twelve years experience in planning and directing system-engineering efforts related to modeling & simulation and their integration with military command and control (C2) systems. Provides on-site technical support in planning for and conducting warfighting exercises and experiments, prototype development, and demonstration of advanced technologies for next generation C2 Systems and Command Centers. He holds a Master of Science degree in Educational Technology.

Ke-Thia Yao is a research scientist in the Distributed Scalable Systems Division of the University of Southern California Information Sciences Institute. Currently, he is working on the JESPP project, which has the goal of supporting very large-scale distributed military simulation involving millions of entities. Within the JESPP project he is developing a suite of monitoring/logging/analysis tools to help users better understand the computational and behavioral properties of large-scale simulations. He received his B.S. degree in EECS from UC Berkeley, and his M.S. and Ph.D. degrees in Computer Science from Rutgers University.

Developing Situation Awareness Metrics in a Synthetic Battlespace Environment

Jacqueline M. Curiel, Michael D. Anhalt
Alion Science and Technology
Fairfax, Virginia
{jcuriel,manhalt}@alionscience.com

John J. Tran, Ke-Thia Yao
Information Sciences Institute/USC
Marina del Rey, California
{jtran,kyao}@ISI.EDU

INTRODUCTION

Problem Description

The "one-the-fly" nature of large-scale human-in-the-loop (HITL) experiments, such as those supported by the Joint Semi Automated Forces (JSAF) simulation federation, mirrors that of actual warfare. The scenarios played out in these types of experiments reflect the continuous interaction among forces (i.e., friendly, hostile, and neutral) over the time course of the experiment so that the situation is dynamic, unfolding over time. These aspects of HITL experiments constrain both the players' capabilities of maintaining accurate Situation Awareness (SA) and the evaluators' attempts to assess players' SA in an effective and timely manner.

The problems associated with assessing SA indicate an interest in further understanding the processes involved in situation awareness during these types of experiments and the continued development of performance metrics. Currently, in HITL experiments, players use sensors to detect the presence of entities and their location, which is necessary for situation awareness but not complete. Additionally, SA depends on identifying the proper context of the experiment. This paper presents our current efforts to develop SA metrics.

Motivation

The motivation for this paper is twofold: 1. previous HITL experiments have yielded a wealth of information that is readily available and, for our purposes, provide a useful base to develop our metrics and 2. current methods of evaluating SA in these types of experiments include observations of players during the exercise and players' reports afterwards. Both measures tend to be subjective, making it more difficult to identify and break down different aspects of situation awareness. We believe that incorporating what we know about situation awareness and situation models with the existing data will help us develop metrics that will help us better understand players' situation awareness.

SITUATION AWARENESS AND SITUATION MODELS

The numerous uses of situation(al) awareness underscore its popularity in research applications. Situation awareness includes an awareness of friendly and enemy troop positions at a specific point in time (Pew & Mavor, Eds. 1998). Another more specific view of SA, Endsley's (1998) three level approach has enjoyed widespread acceptance and has been used in numerous research endeavors to investigate SA. Of interest here is its use in evaluating player performance.

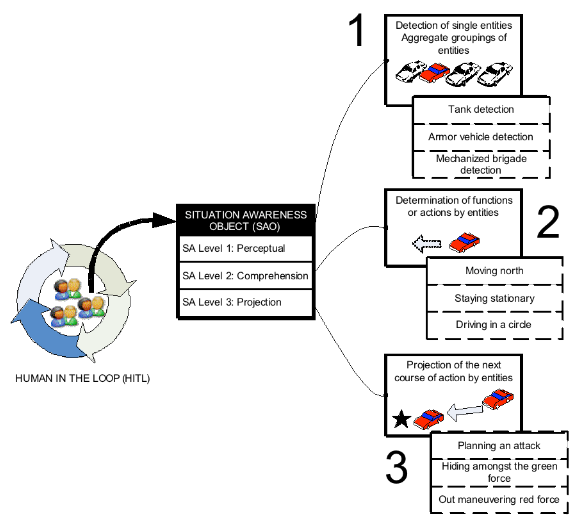


Figure 1. Endsley's SA model specific to synthetic battlespace environment

According to Endsley, SA can be described as three interdependent levels corresponding to: (1) Perceptual SA, (2) Comprehension SA, and (3) Projection SA (Figure 1). The perceptual level involves the detection, recognition, and identification of elements that define a specific situation. Perceptual SA relies on available sensory information, (e.g., from sensors in the case of a player in a HITL experiment) and the player's prior knowledge (e.g., object patterns/schemas activated in memory) to identify individual situation elements and object groups, based on their characteristics.

Comprehension SA reflects an understanding of the current state of affairs and involves making inferences about activities in the current situation. As such, the comprehension level maps the products of perception to object functions. Finally, projection SA consists of interpretations about the trajectory of the situation based on the products of Comprehension SA and prior knowledge. These interpretations include identifying the range of possible trajectories or courses of action along with determining the likelihood of occurrence of each.

At all SA levels is affected by uncertainty due to a number of factors, such as limitations of sensors, and limitations in player's prior knowledge, and the goals of the enemy. Figure 2 shows questions that are relevant to all three SA levels.

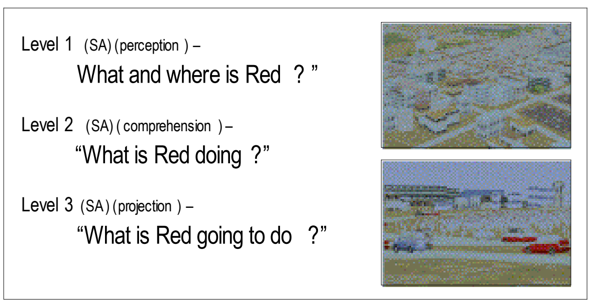


Figure 2. Desired SA Level Metrics in JUO

In our 2004 I/ITSEC paper [1866, Tran, Curiel & Yao], we proposed that the findings of reading comprehension experiments used to study situation models could guide the evaluation of situation awareness in JSAF HITL experiments. Situation Models, mental representations of a situation, are analogous to the mental products of Comprehension SA. Likewise, these representations also depend on the products of lower levels of processing (e.g., textbase and propositional representations in the case of reading comprehension) as well as prior knowledge (e.g., situation schemas).

Zwaan and Radvansky's Event-Indexing model, have focused on providing empirical support for the idea that situation models are multi-dimensional. Although it is unclear how many dimensions can be involved, influences of space, time, entity/protagonist, causality, and intentionality have been observed (e.g., Zwaan & Radvansky). The findings have been interpreted as indicating that readers construct situation models that are defined by these dimensions and updated when changes in the situation occur. Once the story has ended, readers have encoded a completed situation model that is analogous to the "global static summary," an analysis of the end result of the HITL experiment in which the effectiveness of the strategy, the goals of the mission, and the effectiveness of the information provided by the sensors are evaluated. This paper focuses on the relationship between situation awareness and situation model dimensions in HITL a experiment.

EXPERIMENTAL BACKGROUND

Our focus is on the first phase of the Joint Urban Operations (JUO) Urban Resolve experiment conducted by the USJFCOM J9 Directorate and Joint Advance Warfighting Program (JAWP) to guide the development of future sensor capabilities that help soldiers fight in complex urban environments (Ceranowicz & Torpey, 2004). Urban Resolve Phase 1 focused on evaluating the use of human and advance intelligence, surveillance, and reconnaissance technologies to gain situation awareness. Future phases will focus on evaluating the ability to precisely shape the urban battlespace using advanced concept of operations.

Urban Terrain JUO Urban Resolve uses detailed high-fidelity entity-based simulations of urban city areas to exercise proposed sensor capabilities. The Urban Resolve terrain database includes dense urban road networks with over 1.8 million buildings (Prager et al., 2004). Some of these buildings are based on actual real world building footprints, and some have interiors to model parking garages. The terrain features includes elements like parked cars, dumpsters, jersey barriers, individual trees, tree canopies and trashcans. The terrain landscape ranges from deep urban canyons with tall buildings to flat parking areas and open spaces.

This urban terrain is inhabited by approximately 100,000 clutter entities (Speicher & Wilbert, 2004, Williams & Tran 2003). These clutter entities can range from ground vehicles to pedestrians to air/sea vehicles. At the individual entity scale, the ground vehicles follow traffic rules and behave properly at road intersections. At the aggregate scale, the ground vehicles follow the normal flow a bustling city. Rush hours occur during the morning and late afternoon as entities go to and from work. During the lunch hour people go on errand runs, and during the evening people go to restaurants.

Red Force Hiding within this urban terrain is the Red Force (Haskell et al., 2004). The Red Force primarily consists of dismounted infantries, but they also include heavy crew-served weapons, "technicals" (vehicles armed with heavy weapons), light transportation trucks, short-range air defense forces and artillery support. The Red Force follows Techniques, Tactics and Procedures (TTPs). The Red Force tried to blend in the urban environment by pretending to be part of the civilian clutter population, moving about the city to set up fighting locations by fortifying builds and creating booby traps.

Blue Force The objective of the Blue Force is to gain situation awareness of the Red Force. For UR Phase I, the Blue Force is made up of only sensors. The sensors include unmanned aerial vehicles (UAV), low flying organic aerial vehicles (OAV), unattended ground sensors (UGS) and human intelligence. The job for the Blue Cell human players is to task these sensors. observe and track the Red Force. Each Blue Cell human player is given access to a JSAF graphical map display. The map contains the detailed urban terrain overlaid with the positions of the sensors and the Red Force tracks generated by the sensors. The sensors are not completely accurate. The tracks may misclassify the Red entities, and the perceived entity location/velocity may vary from the actual location/velocity.

Procedure

Data for our analyses was obtained from the Urban Resolve Phase 1 set of experiments, which explored new approaches to urban combat. The general procedure follows below.

Participants The Blue Team was comprised of nine active- and reserve-duty military personnel, along with retired military and other contractors. They were selected for the experiment based on previous intelligence experience, or their command and control background, as well as for their ability to adapt to and use new software applications.

Pre-Experimental Training The blue team was given several weeks of training to enable them to become more familiar with application operations, such as the JSAF simulation system and IWS (Information Work Station), to provide briefings about projected enemy capabilities and their likely courses of action, and to provide intelligence briefings to help the players understand their dynamic activities.

Method The Blue Team occupied a room with computers and projected displays. Their main objective was to use their futuristic 2018 sensors to gain situation awareness of the Red Force by controlling their sensor placement and moving them as necessary to follow or anticipate enemy movement.

Each player operated a command and control suite, made up of the JSAF simulation system, with two monitors that displayed a map and allowed for simulation control. They also used a collaborative tool application named Information Work Station (IWS) for chat, email, document sharing and discussions. During the trials, players communicated using Situational Awareness Objects, which recorded players situation information about the enemy, shared map overlays, Voice over Internet Protocol (VoIP), NetTalk chat, and limited face-to-face communications.

The players were briefed prior to each trial regarding enemy capabilities, activities and their likely courses of action. They were told what their sensor limitations were, based on the trial conditions and briefed on any modifications to the JSAF software that might affect their play. The team was flexible in establishing each member's responsibilities and over time, the team decided to have a Commander, with a Sensor Manager and a Surveillance Manger working directly for him. Six Sensor operators worked directly for the Sensor Manager, making sensor asset requests to the Sensor Manager.

Experimental Trials Along with the baseline trial, there were six experimental trials as can be seen in table 1. The type, numbers and capability of the sensors were modified for each trial to determine the impact of the specific changes in the resultant SA. Each trial lasted four or five days, with game play lasting about 7 hours.

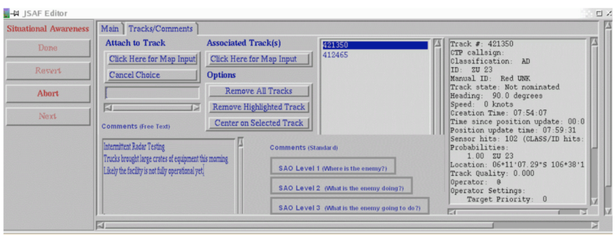
Table 1. Experimental Trials

Trial	Conditions	Duration (hours)
1	Base Case	48
2a	Inactive Red	24
2b	Poor Weather & Inactive Red	18
3a	Signature Reduction	24
3b	½ Inventory	24
4a	No Tags	24
4b	No Tags, ½ Inventory, Signature Reduction	24

For Trial 1, which served as the base case scenario, players had full use of all sensors and the enemy was on the move. For Trial 2a, the enemy moved less frequently and therefore had less exposure to the sensors. For Trial 2b, cloud cover obscured the high altitude sensors and so that there was less initial detection. For Trial 3a, the enemy was allowed to use camouflage. For Trial 3b, the number of sensors was reduced by half. For Trial 4a, futuristic radio frequency (RF) tagging of vehicles and humans was not used, and the enemy could not use camouflage. For Trial 4b, the sensor inventory was reduced by half, the enemy could use camouflage and players could not use RF tagging.

SAO Objects

Our data were obtained from Situation Awareness Objects (SAOs), a method of recording information about red force entities that has only been used this series of experiments. The SAO is a compact package of information that players create and place on a shared terrain map that contains their thoughts, assumptions, and their understanding regarding the enemy. The SAOs are created by selecting options from pull down menus tailored for the trial and modified as the players requested more options. The SAO includes an option to let the players include free-text. Figure 3 shows the SAO screens and sample comments.



Sample Blue standard comments (by SAO category as prescribed by the user)

Level 1	Level 2	Level 3
Appears to be a mortar site	Appears to be a group of insurgents placing a mortar pad	Expect a launch at sunset
Enemy Recon team on roof	Recon team is waiting until street clears	Recon Team will radio mortar team when all clear
Vehicles and Explosives in bldg	Explosives are being loaded into vehicles	Vehicles likely to move to FOB 3 during the night

Sample Red standard comments (by SAO category as prescribed by the user)

Level 1	Level 2	Level 3
Establishing a mortar site	Insurgents placing a mortar pad	Plan a launch at sunset
Positioning a Recon team on roof	Recon team is waiting until street clears	Recon Team will radio mortar team when all clear
Moving Vehicles and Explosives into bldg	Explosives are being loaded into vehicles	Vehicles will move to FOB 3 during the night

Figure 3. SAO input screen

SAOs allow players to quickly enter relevant SA data during the experiment and are shared among other players dynamically and instaneously amongst all the players. They support two complementary objectives: team collaboration and data collection for after action review and data analysis. SAO options are designed to be comprehensive, but not to have players decide the level of SA they refer.

The use of SAOs supplement existing techniques used to assess situation awareness and reduce the analyst's need to intrude on the player's activities in order to assess their performance.

Figure 4 shows that the SAOs are tailored to provide players with relevant real-time data to support their understanding and assessment of the player's SA. Further, SAOs can be used for more in-depth analysis after the experiment trial, allowing analysts to compare actual enemy activities with the SAOs. The SAO approach is successful because the players gain benefit from using SAOs allowing them to share information rapidly and SAOs provide a resource for the analysts to easily and rapidly assess player SA.

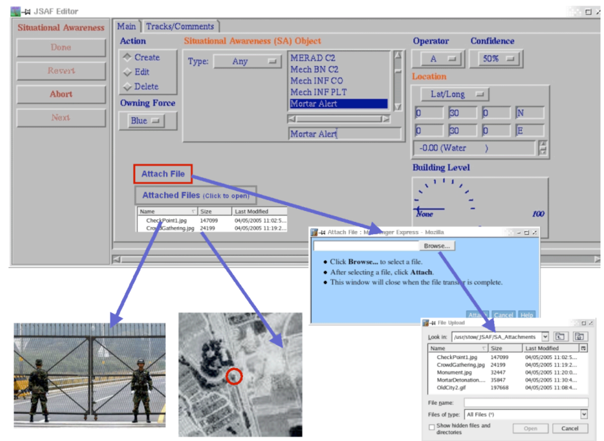


Figure 4. SAOs map to real life model

ANALYSIS

Figure 5 shows counts of the SAO comments for each trial. As can be seen, the Baseline Trial showed the most SAO comments, which is not surprising, given that the duration of that trial was at least twice as long as the other six trials. Of note is that in Trial 4b, which did not use the futuristic RF tags and had both ½ inventory and signature reduction, showed slightly more SAO comments than either the signature reduction trial (3a) or the ½ inventory trial (3b).

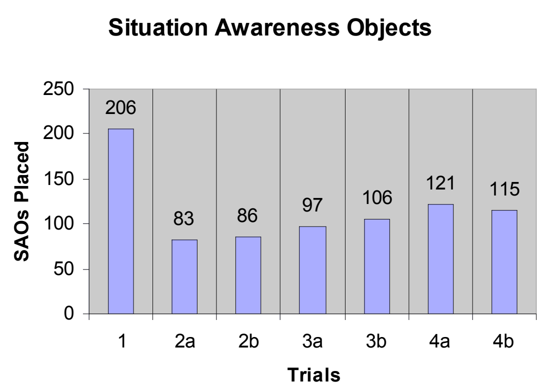


Figure 5. SAO count across the JUO-UR1 trials

Table 2 summarizes the SAO comments along the five situation dimensions and three situation awareness levels (perception 1A and B, comprehension, and projection). The SAO comments were categorized by a trained judge and independently verified. As an example, the comment, "Tank PLT 0755 - Tank platoon heading South from the airport," has three dimensional markers: (i.e., Entity – "Tank Platoon"; Time – 0755; Space – "heading South from the airport") and a first and second level SA (i.e., knowing that Red is a tank platoon and that Red is heading South from the airport). We also make a distinction between SA Level 1A and 1B. For the previous example, the SAO notes a "Tank Platoon," which is identified as level 1B because it is grouped (Platoon). The rest of our analyses are based on these counts.

Effects between the first and second week of trials 2, 3, and 4

As can be seen in Figure 6, there is a decrease in Level 2 SA from the first and second week, for trials 2a/2b, 3a/3b, and 4a/b. This may indicate that the "b" conditions are generally more difficult to identify than the "a" conditions. For example, in trial 2b, in addition to having Red being inactive, there is the additional factor of poor weather that players must contend with. In trial 3, a reduction of Red inventory seems to have a greater effect on Level 2 SA than signature reduction. Finally, in trial 4, the combination of no RF tags, ½ inventory, and signature reduction have a greater effect on Level 2 SA than no RF tags alone.

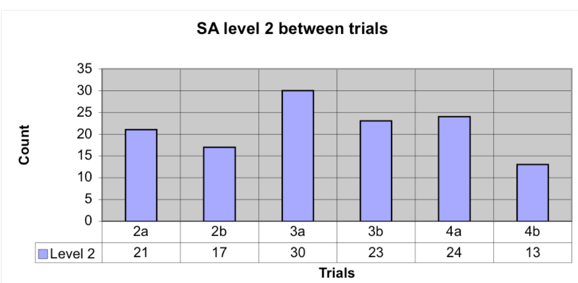


Figure 6. Level 2 SA for each experimental trial

Lower level SA and Situation Dimensions

In looking at Figure 7, it is apparent that entity information, followed by space and time, dominates the SAO comments. In contrast, there are relatively few comments that contain goal and causal information. The sheer amount of entity information reflects the fact the entity information was mentioned in almost every SAO. Additionally, spatial information tended to co-occur with temporal information.

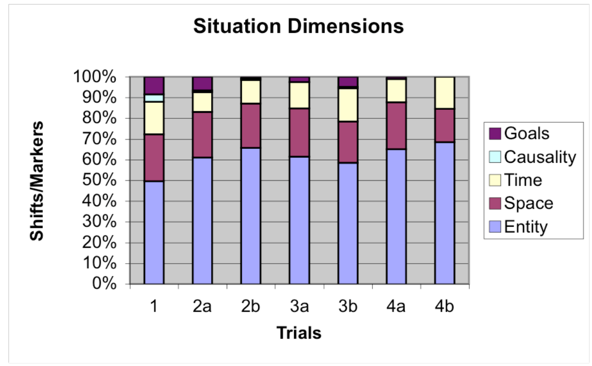


Figure 7. Situation dimensions for the seven trials

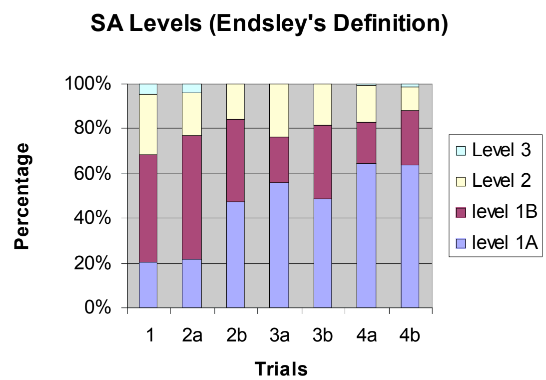


Figure 8. SA levels for the seven trials

Figure 8 shows that Level 1 SA. Similarly, levels 1a & 1b acounts for more than half of the SA levels recorded, and similar level 3 is only a small percentage of the total SA recorded, Figure 8.

Comparison Between Situation Dimensions and SA Levels

Next, we directly compare the situation model dimension counts across the SA Levels. We break this down in the following three figures 9-11. Figure 9 shows that SAOs that refer to entity information tend to be those that include SA information at the Level 1 perceptual level. Figure 10 shows that SAOs that include spatial and temporal information tend to be those that include SA information at the Level 2 comprehension level. Figure 11 did not provide clear cut evidence for a relationship between causal/goal information and SA information at the Level 3 projection level. However, we suspect that this is due to the fact that there are too few data points to make this a reliable comparison. Evidence for a relationship between situation model dimensions and levels of SA implies that efforts to automate situation awareness may consider the information provided by situation dimensions.

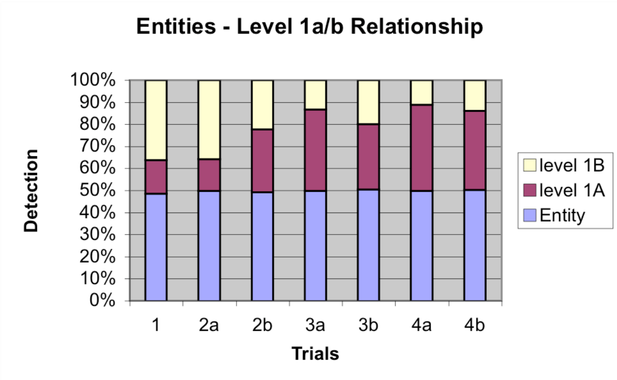


Figure 9. Entities – SA Level 1 Relationship

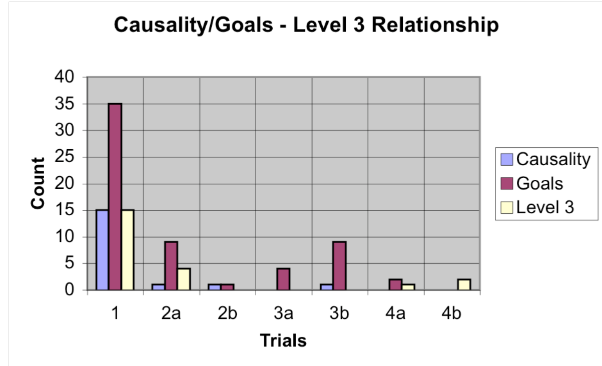


Figure 11. Causality/Goals – SA Level 3 Relationship

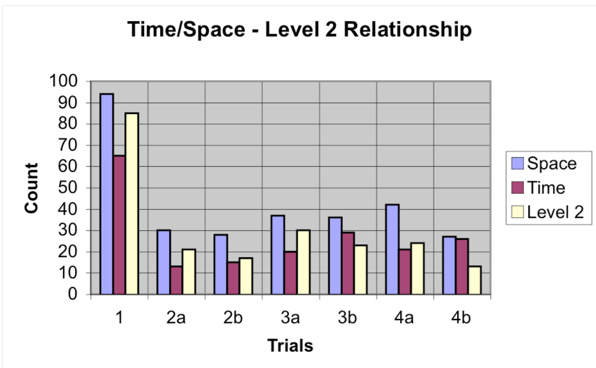


Figure 10. Time/Space – SA Level 2 Relationship

Table 2. SAO data collected for the seven trials of JUO

TRIAL	Entity	Space	Time	Causality	Goals	Level 1A	Level 1B	Level 2	Level 3
1	206	94	65	15	35	65	154	85	15
2a	83	30	13	1	9	24	60	21	4
2b	86	28	15	1	1	50	39	17	0
3a	97	37	20	0	4	72	26	30	0
3b	106	36	29	1	9	62	42	23	0
4a	121	42	21	0	2	95	27	24	1
4b	115	27	26	0	0	82	32	13	2

SUMMARY AND CONCLUSION

In summary, an analysis of SAOs recorded during a JUO Urban Resolve HITL experiment found evidence for a correspondence between levels of situation awareness and the situation model dimensions. Specifically, Level 1 SA comments included a relatively high proportion of spatial and temporal information, whereas Level 3 comments included information about the Red Force goals and intent. Our analysis is also consistent with previous observations that there tends to be more relatively information available about lower levels of SA.

This analysis yielded some interesting observations. Notably, causal information was lacking in players' comments. It is possible that players either did not ascribe causal relationships between events or they did notice causal relationships but did not record them. Determining causality is inherently more difficult than tracking entity locations and may have subsequently been less of a focus for the players. It does seem that increasing the ability to detect causal relationships between events would increase players' situation awareness.

Our approach differs from previous attempts to assess situation awareness in that it is based on entries players made during the experiment, rather than on observations of the players' activities both during and after the experiment. We believe that our approach is advantageous in that it has the potential to allow players to track their situation awareness online.

Future work will focus on addressing this possibility as well as modifying the manner in which data is recorded so that it is done more automatically. We are also interested in comparing our metrics of the players's SA against other methods that capture and analyze simulation groundtruth, e.g. the FAARS's data-collection effort (Graebener 2003) or the Cognitive Enabled ARCHitectures (CEARCH) project.

ACKNOWLEDGEMENTS

We wish to recognize our birddog, LTC Scott Shutzmeister who oversaw the completion of this paper. Special thanks go to Maston Gray, Jim Blank, Rae Denhke, Dr. Robert F. Lucas, Dan Davis, and Donna Brooks-Denhke for their support during various stages of this project. General Larry Budge (USA Ret.), Dr. Gabiel Radvansky, Dr. Andew Ceranowicz provided relevant comments on earlier versions of this paper. And lastly, special thanks go to NuGai Tran for providing the impetus to initiate and complete this project. AFRL Contracts F30602-02-C-0213 and FA875 0-05-2-0204.

REFERENCES

CEARCH. Cognitive Enable ARCHitures. Retrieved 24 June 2005 from http://cearch.east.isi.edu/

Ceranowicz, A. & Torpey, M. (2004). Adapting to Urban Warfare. Paper presented at the Interservice/Industry Training, Simulation, and Education Conference.

Endsley, M. R. (1998). Theoretical Underpinnings of Situation Awareness: A Critical Review. In M. R. Endsley & D. J. Garland (Eds.) Situational Awareness Analysis and Measurement. Mahwah, MJ: Lawrence Erlbaum.

Graebener, R., Rafuse, G., Miller, R. & Yao, K-T. (2003). The Road to Successful Joint Experimentation Starts at the Data Collection Trail. Paper presented at the Interservice/Industry Training, Simulation, and Education Conference.

Lutz, M. F., & Radvansky, G. A. (1997). The fate of completed goal information in narrative comprehension. Journal of Memory and Language, 36, 293-310.

Pew, R. W. & Mavor, A. S. (1998) (eds). Modeling human and organizational behavior. Washington D. C.; National Academy Press.

Prager, S., Cauble, K., Bakeman, D., Haes, S. & Goodman, G. (2004). Malls, Sprawl and Clutter: Realistic Terrain for Simulation of JUO. Interservice/Industry, Simulation, and Education Conference.

Speicher, D., & Wilbert, D. (2004). Simulating Urban Traffic in Support of the Joint Urban Operations Experiment. Interservice/Industry, Simulation, and Education Conference (I/ITSEC).

Speer, N. K. and Zacks, J. M. (2005). Temporal changes as event boundaries: Processing and memory consequences of narrative time shifts. Journal of Memory and Language, (additional info.).

Tran, J. J. Curiel, J. M. & Yao, K. (2004). An Interdisciplinary Approach to the Study of Battlefield Simulation Systems. Paper presented at the Interservice/Industry Training, Simulation and Education Conference.

Willams, R. & Tran, J.J. (2003). Supporting Distributed Simulation on Scalable Parallel Processor System. Paper presnted at the Interservice/Industry Training, Simulation and Education Conference.

Zwaan, R. A., and G.A. Radvansky, G. A. (1998). Situation models in language comprehension and memory. Psychological Bulletin, 123, 162-185.

Columns on last page should be as close as possible to equal length.