ADVANCED DECISION ARCHITECTURES FOR THE WARFIGHTER:
FOUNDATIONS AND TECHNOLOGY

EDITED BY PATRICIA MCDERMOTT AND LAUREL ALLENDER
ADVANCED DECISION ARCHITECTURES
FOR THE WARFIGHTER:
FOUNDATIONS AND TECHNOLOGY

EDITED BY PATRICIA MCDERMOTT
AND LAUREL ALLENDER

This book, produced by the Partners of the
Army Research Laboratory Advanced Decision Architectures
Collaborative Technology Alliance, also serves as the
SECTION III

ACTING ON BATTLEFIELD INFORMATION
Chapter 14

MODELING AND MEASURING SITUATION AWARENESS IN INDIVIDUALS AND TEAMS

CLEOTILDE GONZALEZ, PH.D.
Dynamic Decision Making Laboratory
Carnegie Mellon University, Pittsburgh, PA

LELYN SANER, PH.D.
Dynamic Decision Making Laboratory
Carnegie Mellon University, Pittsburgh, PA

MICA ENDSLEY, PH.D.
SA Technologies, Marietta, GA

CHERYL A. BOLSTAD, PH.D.
SA Technologies, Marietta, GA

HAYDEE M. CUEVAS, PH.D.
SA Technologies, Marietta, GA

INTRODUCTION
Situation Awareness (SA) is a complex construct that cannot be fully understood from a single perspective. Rather, SA entails a multifaceted process in which individual and team factors, at both the micro and macro levels, need to be integrated. For example, research on team SA needs to take into account the factors that contribute to a given person’s individual SA while integrating the factors that contribute to any two team members’ shared SA, as well as team or organizational SA. Also, the complexity of SA arises from multiple individual cognitive abilities including learning, working memory, different levels of perception, understanding, projection, etc. These individual factors interact with multiple team and organizational factors such as geographical distribution, collaborative tool usage, network proximity, similarity of the individuals’ background experiences, and familiarity with each others’ skills, among others.

In this chapter we bring together different levels and factors of SA complexity into an integrated framework. Throughout the duration of the Army Research Laboratories Advanced Decisions Architectures (ARL-ADA) research program, we have addressed the complexity of SA by advancing the development and validity of measures at the individual and team levels. We have also created computational models of SA that have moved forward the descriptive nature of SA into more concrete and formal representations. This chapter summarizes all the past work we have done within this program and attempts to integrate all the previous findings into a research framework. The details of each of these research pieces can be found in the previous original publications. The integrated framework presented here came from a combination of scientific research methods at the micro and macro levels that included computational models, and laboratory and field studies. The framework integrates our past findings from experimental laboratory studies and ACT-R cognitive computational models on individual SA to field experiments and social network analysis of team and shared SA.

**COMPUTATIONAL MODELS OF SA**

Computational models have been developed to address individual cognitive aspects of SA, such as recognition, perception, and learning; to address design aspects of SA, such as visual interfaces and goal directed task analysis; and to address organizational and team aspects of shared SA, such as communication network distance, physical proximity, and knowledge and background similarity. The existence of these different models illustrates the challenge that one faces in the study of SA generally. Efforts demonstrate a positive approach to advance our scientific understanding of SA because they address aspects of the problem that are salient from different perspectives. This research demonstrates that
individual and organizational levels of SA are not incompatible, but rather they complement each other. As suggested recently (Lebiere, Gonzalez, & Warwick, 2009), because much of the history of computational cognitive modeling has been microcognitive, these psychological architectures are a good starting point to develop macrocognitive or organizational models of SA.

**Computational Models of Individual Aspects of SA**

Many computational models have been developed to address some important aspects of SA at the individual level, such as perception, attention, memory, and learning among others (Gonzalez, Juarez, & Graham, 2004; Juarez & Gonzalez, 2003, 2004; McCarley, Wickens, Goh, & Horrey, 2002). Some of these models have been framed in the context of military applications such as perceiving and understanding a situation in the battlefield, and many of these models involve leading cognitive architectures such as ACT-R (Anderson & Lebiere, 1998) and SOAR (Newell, 1990). Here we summarize our own development of cognitive models of situation awareness (CMSA) at the individual level by using the ACT-R architecture (Anderson & Lebiere, 1998).

A SA meta-architecture was developed to propose a conceptual design for CMSA (Gonzalez et al., 2004; Juarez & Gonzalez, 2003). The meta-architecture presented a set of modules that encapsulate the most essential cognitive aspects of SA. A demonstration of the architecture was developed in a commanding decision making mission in the OneSAF Test Bed (OTB), which interacts in real-time with the ACT-R cognitive model. This architecture consists of a recognition, assessment, prediction, and control modules. The recognition module gathers and encodes visual sensory information from the environment (in OTB) and encodes the information into ACT-R representations. The assessment module represents and manipulates information and updates, and maintains information on the environment (e.g., updates information about location, status, and types of entities in the battlefield). The prediction module constructs and evaluates hypotheses about the function of an entity in a plan to determine the probability and possible success of a course of action. The control module then provides the capability to select actions that will change the state of the world.

The SA meta-architecture is used as an experimental platform to study the relationships among the model parameters and SA performance (Juarez & Gonzalez, 2004). The architecture allows one to manipulate a number of parameters and evaluate their impact in SA, including parameters of the task such as scenario complexity and the 'human' (ACT-R model) parameters like experience.
A major question arising from these developments was which method to select for evaluating the SA of the model. One widely tested and validated approach to assessing SA in human subjects is the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1995a). SAGAT allows for immediate assessment of SA by querying operators on their current perceptions, assessments, and projections of the situation (SA levels 1, 2, and 3 respectively). SAGAT has been empirically validated with regard to its utility to provide valid and reliable assessment of SA across a variety of domains (Endsley, 1990; Endsley, Sollenberger, & Stein, 1999; Hogg, Torralba, & Volden, 1993; Matthews, Pleban, Endsley, & Strater, 2000; Riley & Kaber, 2001). As with human SA measurement, we used SAGAT to measure SA for our computational models. The validation with SAGAT was implemented successfully (Juarez & Gonzalez, 2004); however, the cognitive aspects of the model still need to be validated with human data. A validation of the model’s predictions requires the comparison of SAGAT data from human participants to the model’s SAGAT data. Obtaining realistic human data in this context continues to be a challenge for SA modeling research. However, cognitive models have many other practical applications. One of them, presented next, is the use of computational models in conjunction with simulation tools in the evaluation of graphical interfaces.

Computational Models of Design Aspects of SA

Because of the relevance of SA in Army operations, methods and tools that help us predict and mitigate low levels of SA are valuable for researchers and interface designers alike. In fact, accurately and reliably assessing how well an interface design concept supports user SA is essential to any design process. Given the limited time and funding available for developing multiple high fidelity prototypes, designers need a tool that will allow them to evaluate multiple design concepts early in the design cycle to minimize expense while maximizing the ultimate utility of the user interface.

The ability to quickly predict the SA afforded by a particular display design is essential for the success of the computer systems researchers develop. Our research described in Gonzalez, Juarez, Endsley, and Jones (2006) presents how the computational architecture and ACT-R model of SA were used in a new application for the prediction of the SA elicited through a graphical user interface (GUI), and how the CMSA could be used in a new tool for prototyping and evaluating the user’s SA through simple GUIs. This work is summarized next.
The SA-oriented design process (Endsley, 2003; Endsley, Bolte, & Jones, 2003), provides a key methodology for developing user-centered displays by focusing on optimizing situation awareness. By creating designs that enhance an operator’s awareness of what is happening in a given situation, SA can improve dramatically. The design process starts with the gathering of SA requirements. These are determined through a cognitive task analysis technique called Goal-Directed Cognitive Task Analysis (GDTA). A GDTA identifies the major goals and sub-goals for each job; the critical decisions the individual must make to achieve each goal and sub-goal; and the situation awareness requirements needed for making these decisions and carrying out each goal.

These SA requirements focus not only on what data an individual needs, but also on how that information is integrated or combined to address each decision. This process forms the basis for determining the exact information (at all three levels of SA: perception, comprehension, and projection) that needs to be included in the display visualizations. A final step of the SA-oriented design process emphasizes the objective measurement of SA during man-in-the-loop simulation testing using SAGAT. SAGAT provides a sensitive and diagnostic measure of SA that can be used to evaluate new interface technologies, display concepts, sensor suites, and training programs (Endsley, 1995; Endsley & Garland, 2000).

The Designer’s Situation Awareness Toolbox (DeSAT) was created to assist designers in carrying out the SA-oriented design process (Endsley, 2005). It includes (1) a software tool for easily creating, editing, and storing effective GDTAs, (2) A GDTA checklist tool, to aid designers in evaluating the degree to which a display design meets the SA requirements of the user, (3) A SA-oriented design guidelines tool, which guides designers in determining how well a given design will support user SA, and (4) A SAGAT tool, which allows designers to rapidly customize SAGAT queries to the relevant user domain and to administer SAGAT during user testing to empirically evaluate display designs.

The concept of DeSAT was used as a starting point to develop a new CMSA tool for predicting and evaluating how different GUI designs elicit SA. The CMSA tool (described in detail in Gonzalez et al., 2006) integrates the GDTA process through the DeSAT to generate a CMSA in a semi-automated way. The process of generating the CMSA from the GDTA is presented in Figure 14.1. The GDTA, which is the DeSAT output (an .xml file), and the image files that define a GUI are the two starting points of the process. The GDTA description consists of a list of goals, decisions, and information requirements to support situation awareness in a very specific domain (Endsley, 2005). The GUI is described in terms of its graphical components including widgets (e.g., buttons, graphical sections, icons) and their behaviors, as well as their organization in the graphical interface. The tool also allows one to define the domain knowledge and to identify the background knowledge needed to perform a task.
As a next step, the tool defines correspondence mappings between the GDTA components and specific widgets or graphical elements defined in the GUI. The mapping process helps produce a high-level definition language for SA from which an ACT-R model of SA is then generated and executed through the ACT-R architecture (Anderson & Lebiere, 1998).
The functioning of this CMSA tool was demonstrated with a realistic example of a GUI prototype design (Gonzalez et al., 2006). The example used a selected set of GUIs to perform the logistics tasks of a company unit (see example in Figure 14.2). The user of this system was expected to plan supplies, project the demand for a specific time period, and obtain the consumer trends for some items during specific time periods. The task involves high-level cognitive processes that sometimes are not necessarily expressed by a click of a button in a GUI. This logistics task consists of four sub-goals, each of which is represented in one of five screens in the application.

![Figure 14.2](image_url)

**Figure 14.2**
Example of a GUI used in the demonstration of the evaluation of the SA in Gonzalez et al. (2006)

In this example, the GDTA defines information requirements, decisions, and goals. Although the GDTA can be interpreted by a human, the computer model needs domain knowledge to understand the GDTA components. This CMSA tool defines domain objects and constraints to support the model on making decisions. Knowledge about domain objects in our example consists of: roads, road condition and road type; terrain type, elevation, type of soil; type of
vehicle, strength, capabilities, location; enemy threats; fuel requirements; mission, supply and shipping schedule; history of consumption, requirements, and usage rate. In addition, the system requires the knowledge about the criteria to select alternatives. For example, the system must contain definitions for best path which is mapped to shortest distance, minimal threats, and road conditions.

The CMSA tool produced a set of commands to generate a memory chunk for every element in the GDTA, for every element in the graphic interface, and for every domain object and domain constraint. Finally, the system generated ACT-SA code with the procedural description of every task. The system then can be executed to collect the performance times and the number of operations needed to perform a task.

This work, described in Gonzalez et al. (2006), demonstrates the practical applications of CMSA. Most of the tools currently known to automatically generate cognitive models from graphical user interfaces have been created to evaluate the interface itself. That is, most of the work done on this area generates a model to evaluate the usability or the human-computer interaction of the interface. In addition, most of the tools developed up to date deal with very simple interface actions, such as point, click, and move. Our tool focuses not on the evaluation of the interface elements, but rather on the evaluation of the SA afforded by an interface. As SA involves high level actions, such as prediction of future status, it is important to include high level operators in our models, beyond simple interface actions.

**Computational Models of Team and Organizational Aspects of SA**

Computational models of SA at the team and organizational levels have also been developed to represent organizational aspects of SA, such as communication, shared understanding, shared workload, etc. Successful models have been based on Social Network Analysis (SNA) theory developed under this ARL-ADA program (Graham, Gonzalez, & Schneider, 2007). This work is summarized here.

SNA is a technique that seeks to quantify the relationships among people in an organization. People and organizations are represented as nodes in a network, and the relationships (e.g., information flows) between people are represented as lines drawn between these nodes. Thus, a social network is a graph consisting of individuals and connections among them, where each connection is associated with some form of communication or relationship between the nodes (Borgatti, 1994).
Communication data among an organization’s members can be gathered from shared e-mail headers, chat room traffic, instant messaging, and phone calls, or by surveying the individuals (Wasserman & Faust, 1994). While each of these communication media has different qualities, measure relevance is determined by the organizational context and collaborative tool characteristics.

Graham, Gonzalez, and Schneider (2007) gathered communication data in a field experiment conducted at the Fort Leavenworth Battle Command Battle Laboratory. The US Army was in the opening phase of a ten-year organizational design process for a knowledge-centric command and control element. In support of this initial effort, the Fort Leavenworth Battle Command Battle Laboratory (BCBL) was conducting the first high fidelity experiment to determine organizational constructs that would support command and control in the Transformation Force. The experiment assumed a network-centric staff cell structure supported by a higher level of automation.

Figure 14.3 from Graham et al. (2007) demonstrates a SNA representation of 90 minutes of communication relationships in a 10-person command-and-control cell of the network organization. The 10 nodes are members of a Command Integration Cell (CIC) designed to coordinate the activities of other functionally oriented cells, and constitute a subset of the 56-member prototype network organization that was the focus of this research.
We used traditional SNA measures such as network density, network distance, physical distance, and self-forming teams to draw conclusions about the shared SA of this team, and to ultimately provide feedback to the Battle Lab on unit configuration.

Conclusions on Computational Models of SA

As summarized above, our past research has produced computational representations of different aspects of SA including cognitive, design, and organizational levels. These research efforts demonstrate the complexity in representing and reproducing human SA at different levels of specificity. These efforts also suggested that the development of SA measures, at both the individual and organizational levels, is essential to make progress in assessing SA, both behaviorally and computationally.

While the SA literature base is large, many of the metrics currently used to measure SA focus only on component aspects of SA and don’t necessarily measure SA as an identifiable, integrated construct. This is partly due to the fact that the definitions of SA and shared SA are still under some debate. The need for better measures is particularly clear with respect to Shared SA. In order to improve the computational representations of SA, researchers need to develop valid approaches to assess both individual and shared SA.

MEASURES OF SA

Over the duration of this research program, we have investigated SA at both the individual and team levels. At the individual level we focused on the learning of SA and the influence of individual aspects of SA using laboratory dynamic decision-making tasks. At the team level, we focused on the organizational measures of SA through the data collected in several field experiments in the Army Laboratories such as Ft. Leavenworth and the Joint Personnel Recovery Agency (JPRA). These efforts are summarized next.
Measures of Individual SA

Task practice may allow people to develop “the mental models, schemas, and goal-directed processing that are critical for SA in most domains” (Endsley et al., 2003). An important first step in designing more effective displays and cognitive models of SA is to understand how individuals naturally improve their SA through practice. Similarly, cognitive mechanisms, such as working memory and learning or experience, are commonly assumed to influence SA (Endsley, 1995; Endsley & Robertson, 2000). Although there is some clear evidence regarding the influence of working memory on SA (Durso, Bleckley, & Dattel, 2006), we only have a weak understanding of the relationship between working memory and experience.

A common method to measure SA described earlier in this chapter is SAGAT. In this method a task simulation is stopped at random points and a participant answers a set of queries about the situation. Queries may be answered while the simulation display is not visible or covered (Endsley, 1995) or while the display is visible and uncovered (Durso et al., 1995). Gonzalez and Wimisberg (2007) conducted a laboratory experiment in which these two individual SA measurement conditions were used while participants played a computer simulation over several days. SA was measured using two different query conditions methods, a covered condition in which queries were asked while the display was blanked out and an uncovered condition in which queries were asked while the display was shown. A working memory measure was collected from participants as well.

Results, reported in detail in Gonzalez and Wimisberg (2007), showed that SA improves with practice when measured in a covered display condition, but not when measured in the uncovered condition. Furthermore, the effect of experience appears only in the covered condition for perception and comprehension queries, not for projection queries. Another main conclusion from this empirical study was that the moderating effect of working memory on SA changes with task practice, and depends on the conditions in which SA is measured. The participant’s level of working memory predicted SA scores only in the covered condition, such that, when the SAGAT requirement of blanking out the screen while an operator answers queries was implemented, working memory became a significant factor for SA. We also found that the effect of working memory was more relevant for SA at the perception level than for SA at the projection level. Because of the nature of perception-level queries, correct answers fully depend on the visibility of the elements on the display. In contrast, accurate answers to projection-level queries depend on a deeper understanding of the task.
Finally, we found a significant decrease in the correlation between SA and working memory with practice in the task. This result suggests that as experience accumulates, there is a decreasing need to use working memory to maintain SA, although working memory may still be necessary to task performance. Please refer to the original publication (Gonzalez and Wimisberg, 2007) for the details of the design and results.

**Measures of Shared SA**

Our work has extended from the individual aspects of SA to SA in teams and organizations (Graham, Gonzalez, & Doyle, 2003; Graham, Gonzalez, & Schneider, 2007; Saner, Bolstad, Gonzalez, & Cuevas, In Press). In order to function effectively as a team, operators need to develop an accurate understanding of the situation, that is, possess a high level of situation awareness (SA). If people are working in teams and any one of the team members has poor SA, it can undermine the success of the entire team. However, it is often the case that not all team members need to know all of the same information in a given situation (Endsley, 1995). In many situations, individuals in a team possess specialized knowledge and they rely on each other to perform particular tasks. As such, although each team member needs to have good individual SA on the information that is relevant to his/her job, the similarity of the individual SA among team members is only important with respect to their shared task requirements (Saner et al., In Press).

Our approach, explained and validated in Saner et al. (In Press), is based on assessing shared SA from objective measures of individual SA. We used the SAGAT methodology and compared participant responses to ground truth (reality) to evaluate the accuracy of their SA at a given moment in time. Despite the general validity of most existing methods for calculating SA scores, one limitation that they still possess is they often fail to account for errors in SA, in terms of both the amount of information that is thought to be relevant and in the accuracy of a person’s knowledge of it. As such, assessments of SA derived from these methods may be inflated. In order to address both accuracy and similarity of SA, we developed a method for calculating shared SA that first derives true SA scores and then assesses the similarity between the scores of two individuals. Our preliminary analysis suggests that failure to compensate for error in SA might lead to overestimation of performance in a situation.

In addition to developing this new measure of shared SA, Saner et al. (In Press) evaluated whether it was related to cognitive and social factors that often influence performance in team contexts. We investigated SA in two interdisciplinary, multi-team military training exercises. The first data set was obtained by measuring SA among 17 participants engaged in a joint rescue
operation training exercise at the Joint Personnel Recovery Agency (JPRA). Participants were divided into four physically distributed team cells (Army, Navy, Special Forces, and Joint Operations), and had to coordinate rescue maneuvers in response to critical events. In the exercise, participants completed several distinct scenarios over several days, and were rotated into different cells for each scenario. The second data set was derived from a battalion command training exercise at Ft. Leavenworth. In this exercise, 24 participants were divided among four battalion level command positions and five core operational groups (Command Integration, Fires and Effect, Information Superiority, Maneuver and Support, and Build and Sustain), which were physically distributed. Participants remained in the same assigned positions throughout the duration of this exercise. In addition, there were clearly defined roles within each group, including officers in charge. The latter exercise was much more complex and there were more groups coordinating and more levels of rank, the situation was much more complex in this exercise.

The initial expectation was that the more distance there is between an operator and the operational center of the organization, the less SA that operator will have. Results from the rescue operation data set revealed a significant relationship between shared SA and participants’ distance from the joint service cell, which was expected to act as the organizational hub of the C3 structure in that context. In contrast to what was expected, shared SA was significantly better the further away participants were from the hub. This finding, though counterintuitive, did provide evidence that an individual’s role and position within an organization affect the level of shared SA that can be achieved with other individuals in the network. One possible explanation is that participants in the peripheral teams were doing more direct processing of information and did not need the help of the joint service cell to utilize information effectively, but this and other possibilities are still under investigation.

The analysis of battalion command data is still in progress, and is focused more specifically on the social network factors. Cognitive load data was again collected, but detailed data on years of experience and communication were unavailable. However, the role of individuals within cells is being added as a predictor. The Command Integration cell was also an organizational hub in this exercise, so participants’ distance from this cell and physical distance from each other are factors in this analysis as well. The goals of this analysis are to replicate the relationships between the predictors and shared SA that are common to both samples, and also to extend our assessment of the social dynamic influences that were observed in the recovery operation context.

Main Conclusions from Measures of Individual and Shared SA
The success of the CMSA depends on finding appropriate and robust measures of individual and shared SA. The work described in this chapter extends the current state of the art on SA modeling and measurement. In addition, our shared SA measure builds upon work conducted with individual SA measures in an effort to develop new ways to account for error in SA that might otherwise misrepresent performance in a situation.

We validated and extended the current understanding of individual measures of SA, developed procedures to measure the degree of shared SA between two team members and to improve the accuracy of shared SA scores, and created computational models that expand individual and team SA. In future work on the measurement of shared SA, we intend to address more of the dynamic factors involved in team work (e.g., information flow, communication flow, physical location, etc.), particularly the specific roles adopted by individuals and the distribution of task experts within the larger team. If a team is ultimately successful in completing their task, there must be a way to describe why particular knowledge, actions, or coordination patterns led to success. Without a measure of what particular beliefs the participants were operating on, and whether participants were operating on similar or different beliefs when they took similar actions, there is no sure way to reproduce success in later, similar situations. The effect of SA on performance is a question that is still unanswered in Shared SA research, and we plan to pursue this in future work.

ACKNOWLEDGMENTS

This work was supported by the Army Research Laboratory (DAAD19-01-2-0009) award to Cleotilde Gonzalez and Mica Endsley. Many others contributed to the research summarized in this paper. We would like to particularly thank Dr. Octavio Juarez for his contributions in building some of the tools and models summarized in this paper and originally reported in respective publications.

REFERENCES

methodologies. In D. J. Garland & M. R. Endsley (Eds.), Experimental analysis and measurement of situation awareness (pp. 295-303). Daytona Beach, FL: Embry-Riddle Aeronautical Press.


