NAVY SMALL BUSINESS INNOVATIVE RESEARCH PROGRAM

AWARD NUMBER: N04-119

CONTRACT NUMBER: N00014-04-M-0155

DATA ITEM NUMBER: 0001AC

TITLE: Automated Communication Analysis for Interactive Situation Awareness Assessment

PRINCIPAL INVESTIGATOR:
Cheryl A. Bolstad, Ph.D.
SA Technologies
3750 Palladian Village Drive
Building 600
Marietta, GA 30066

CONTRACTING ORGANIZATION: Office of Naval Research

SBIR PROGRAM MANAGER: Cathy Nodgaard

REPORT DATE: 11/01/04

TIME PERIOD COVERED: May 03, 2004 – October 31, 2004

TYPE OF REPORT: Phase I Final Report

SCEURITY CLASSIFICATION: Unclassified

DISTRIBUTION STATEMENT:

- [X] Approved for public release; distribution unlimited
- [ ] Distribution limited to U.S. Government agencies only; report contains proprietary information

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 074-0188

1. Agency Use Only (Leave blank)	2. Report Date	3. Report Type and Period Covered (i.e., annual 1 Jun 00 - 31 May 01)
	11/01/04	Phase 1 Final Report May03-Oct31 2004

4. Title and Subtitle

Automated Communication Analysis for Interactive Situation Awareness Assessment

5. Award Number

N04-119

6. Author(s)

Dr. Cheryl A. Bolstad (SA Technologies)
Dr. Cleotilde Gonzalez (CMU)
LTC. John Graham (CMU)

7. Performing Organization Name (Include Name, City, State, Zip Code and Email for Principal Investigator)

SA Technologies, Inc
3750 Palladian Village Drive
Building 600
Marietta, GA 30066
cheryl@satechnologies.com

8. Performing Organization Report Number (Leave Blank)

9. Sponsoring/Monitoring Agency Name and Address

Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217-5660

10. Sponsoring/Monitoring Agency Report Number (Leave Blank)

11. Supplementary Notes (i.e., report contains color photos, report contains appendix in non-print form, etc.)

12a. Distribution/Availability Statement (check one)

X Approved for public release; distribution unlimited
Distribution limited to U.S. Government agencies only - report contains proprietary information

12b. Distribution Code

(Leave Blank)

13. Abstract ***(Maximum 200 Words)*** *(abstract should contain no proprietary or confidential information)*

The SBIR Phase I goals were to (1) to develop and eventually validate a computational model of shared Situation Awareness (SA) and (2) identify a potentially unobtrusive measure of SA for military teams. The following objectives were achieved in Phase I: (1) assess the potential teams and tasks (identify data collection and model validation needs); (2) assess and create the measures used in the model creation (review the modeling literature and select factors that are related to SA, mental models and shared cognition); (3) collect and analyze the data (data was gathered from three separate military exercises); (4) perform the Social Network Analysis (SNA) and create the shared SA model and (5) identify other factors that need to be included in the model to improve the SA prediction capabilities as well as identify model validation needs. Our research strives to create both a model of shared SA as well as an unobtrusive measure of SA. The model may be thought of as a low cost, unobtrusive, real-time measure of SA. By understanding the various factors that make up shared SA we can create a predictive model of SA using these factors as well as a real time measure of SA.

14. Subject Terms (keywords previously assigned to proposal abstract or terms which apply to this award)

Situation awareness, modeling, communication, team performance, team assessment

15. Number of Pages (count all pages including appendices)

53

16. Price Code (Leave Blank)

17. Security Classification of Report	18. Security Classification of this Page	19. Security Classification of Abstract	20. Limitation of Abstract
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

Automated Communication Analysis for Interactive Situation Awareness Assessment

The SBIR Phase I goals were to (1) to develop and eventually validate a computational model of shared Situation Awareness (SA) and (2) identify a potentially unobtrusive measure of SA for military teams. The following objectives were achieved in Phase I: (1) assess the potential teams and tasks (identify data collection and model validation needs); (2) assess and create the measures used in the model creation (review the modeling literature and select factors that are related to SA, mental models and shared cognition); (3) collect and analyze the data (data was gathered from three separate military exercises); (4) perform the Social Network Analysis (SNA) and create the shared SA model and (5) identify other factors that need to be included in the model to improve the SA prediction capabilities as well as identify model validation needs. Our research strives to create both a model of shared SA as well as an unobtrusive measure of SA. The model may be thought of as a low cost, unobtrusive, real-time measure of SA. By understanding the various factors that make up shared SA we can create a predictive model of SA using these factors as well as a real time measure of SA.

Acknowledgments

Special thanks go to Dr. Michael Letsky of the Office of Naval Research, Human Systems Department and Dr. Norm Warner of Naval Air Systems Command, Human Systems Integration Department for their support throughout the Phase I SBIR, Automated Communication Analysis for Interactive Situation Awareness Assessment.

Our team recognizes the support of the Joint Forces Command, Joint Personnel Recovery Agency (JPRA), the Ft. Knox Battle Command Battle Lab (UAMBL) and the Ft. Leavenworth Battle Command Battle Lab for their assistance in conducting the research needed to validate and create our shared Situation Awareness model. We acknowledge the following individuals from the JPRA: Mr. Fred Kleibacker, Mr. Aaron Zaccagnino and Mr. Palmer Triplett, Mr. Gene Satterfield, Mr. Bruce Townssend, Mr. Jon Couch, Major Martin Littlehales RAF (British Army) and Colonel Mark Bracich for generously volunteering their time to our project.

List of Tables

List of Figures

1.0 Introduction

Our Technical Approach Supports NetCentric Warfare

- Distributed command teams
- Large or small C^3I organizations
- Nodal organizations
- Dynamic teams

Understanding the nature of Situation Awareness (SA) poses a considerable challenge to the research community. As the military undergoes significant changes including smaller, more deployable dispersed forces the need to find new methods to analyze and assess team behavior has increased significantly. This need is especially apparent in future asymmetric warfare operations where soldiers will need to capitalize on their strengths and be aware of their own team's abilities and limitations.

In this new modernized military, if soldiers are to function in a distributed fashion they will need similar mental models and a high degree of shared SA to function effectively. Research has shown that when team members possess similar mental models their team performance is enhanced (Stout, Cannon-Bowers, Salas, & Milanovich, 1999) and when teams are not allowed to generate shared mental models they perform significantly worse than teams with shared mental models (Bolstad & Endsley, 1999). Research has shown that communication increases team effectiveness by helping teams form shared SA and shared mental models (Brannick, Roach, & Salas, 1993; Williges, Johnston, & Briggs, 1966).

Figure 1. Navy Operations

The main focus of this phase 1 SBIR is to develop and validate a computational model of Shared SA that is applicable to dynamic organizations. In this research, we used state of the art techniques in cognition, situation awareness theory, Cognitive Task Analysis (CTA) and Social Network Analysis (SNA). Our approach is on the cutting edge of new research methods. It provides a validated technique for determining shared SA in multiple domains as well as a real-time, unobtrusive assessment measure of SA as well.

1.1 Situation Awareness

In order to measure or model SA one needs to have a thorough understanding of the SA construct. Endsley (1995b) formally defines SA as "…the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future " (p. 36). The definition encompasses several concepts that are important in understanding the SA construct. First, SA is comprised of three levels: perception, comprehension and projection. Level 1 SA, perception, involves the sensory detection of significant environmental cues.

Figure 1 illustrates our focus on distributed military teams. Currently, like other the services, the Navy is undergoing a major transformation. For instance, the Fleet Response Plan (FRP) calls for the Navy to have four fewer carriers than the old fleet paradigm. As ADM Vern Clark, chief of Naval Operations told members of the U.S. Chamber of Commerce in Washington, D.C., (December 3, 2003), "I'm convinced that …we can get the job done with fewer people."

Perception is an active process whereby individuals extract salient cues from their environment (Dominguez, 1994). By selectively directing attention to the incoming stimuli, important/essential information is attended to while nonessential items are disregarded. Level 2 SA, comprehension, involves integrating and comprehending the information in working memory (Eduardo Salas, Prince, Baker, & Shrestha, 1995) to understand how the information will impact upon the individual's goals and objectives. This involves combining bits of information together to form a comprehensive picture of the world, or of that portion of the world of concern to the individual. Level 3 SA, projection, involves extrapolating this information forward in time to determine how it will affect future states of the operating environment (Endsley, 1988; 1993). Level 3 SA combines what the individual knows about the current situation with their mental models or schemata of similar events to predict what might happen next.

Additionally, SA has a temporal and locational component. Time is also an important concept in SA, as SA is a dynamic construct, changing at a tempo dictated by the surrounding action. As new inputs enter the system, the individual incorporates them into this mental representation, making changes as necessary in plans and actions in order to achieve the desired goals. SA also involves knowledge about the activities and events occurring in a specific location of interest to the individual. Thus, the concept of SA includes perception, comprehension and projection of situational information, as well as locational and temporal components.

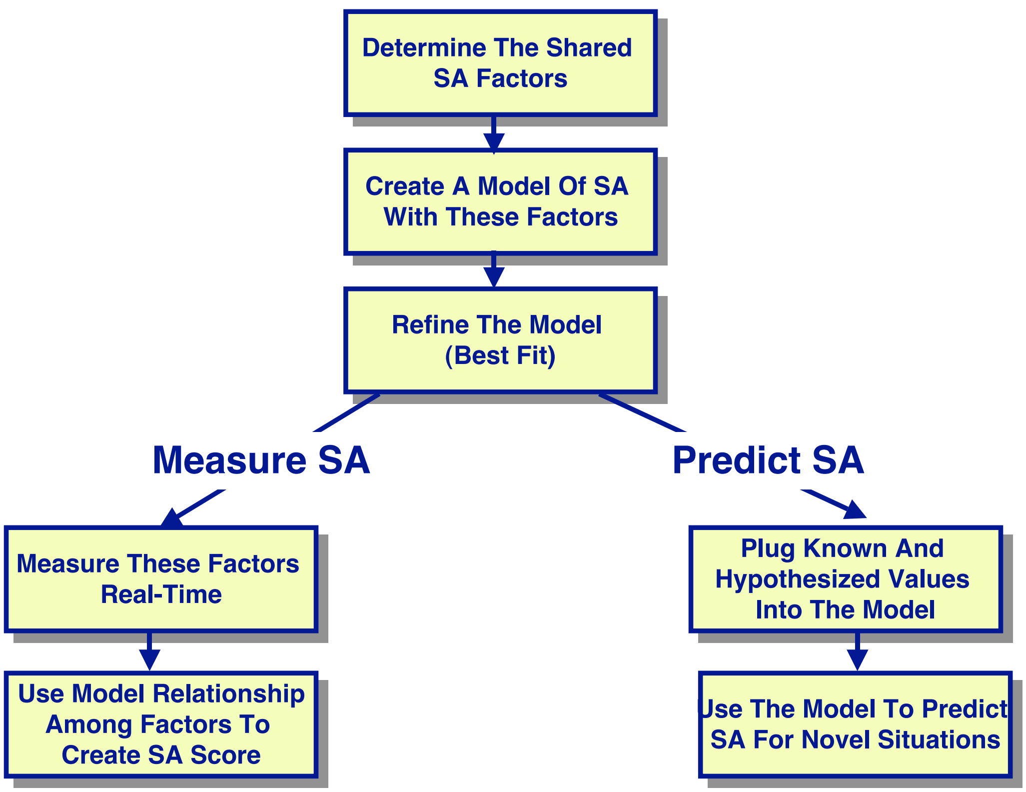


Figure 2: Our Approach

Figure 2 shows the approach we are taking with this research. Our efforts enable us to follow a two-prong approach: measuring SA and predicting SA. We believe that by first gaining an understanding of what variables are needed for SA formation and how these variables work together we will ultimately create more robust and accurate measures of SA. Additionally, our research has indicated that SA is not a simple construct that can be measured with only one variable and therefore this type of approach is warranted.

1.2 Team and Shared SA

Team SA

For this research project we are interested in measuring team and shared SA (not individual SA) as most military environments involve teams of individuals working together towards a common goal. Overall, team SA can be conceived as "the degree to which every team member possesses the SA required for his or her responsibilities" (Endsley, 1995a). Team members need to posses a shared understanding of the situation with regard to their shared SA requirements to develop effective team performance.

SA is also highly related to team goals (Endsley & Jones, 1997). In a team, each crew has a subgoal pertinent to his/her specific role that feeds into the overall team goal. Associated with each team member's subgoal is a set of SA elements about which he or she is concerned. As the members of a team are essentially interdependent in meeting the overall team goal, some overlap between each member's subgoal and their SA requirements will be present. It is this subset of information that constitutes much of the team coordination. If each of two team members needs to know a piece of information, it is not sufficient that one knows it perfectly but the other does not. Instead, each and every team member must have SA for all of his or her own SA requirements or become the proverbial chain's weakest link.

Shared SA

In smoothly functioning teams, each team member shares a common understanding of what is happening in regards to common SA elements. This is known as shared SA:"…the degree to which team members possess the same SA on shared SA requirements" (Endsley and Jones, 1997, p. 54). This principle explains the overlap between the SA requirements of the team members as shown in figure 3. As presented by the clear areas of the figure, not all information needs to be shared. Clearly, each team member is aware of information that is not pertinent to the others on the team. Sharing every detail of each person's job would only create a great deal of noise to sort through to get the needed information.

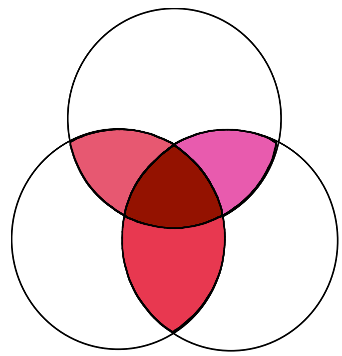


Figure 3. Shared SA Requirements (from Endsley & Jones, 1997, 2000)

Developing shared SA has been hypothesized to involve four factors: (1) Shared SA Requirements – the degree to which team members understand which information is needed by other team members, (2) Shared SA Devices – including communications, shared displays and a shared environment, (3) Shared SA Mechanisms – such as shared mental models, and (4) Shared SA Processes – consisting of effective team processes for sharing relevant information.

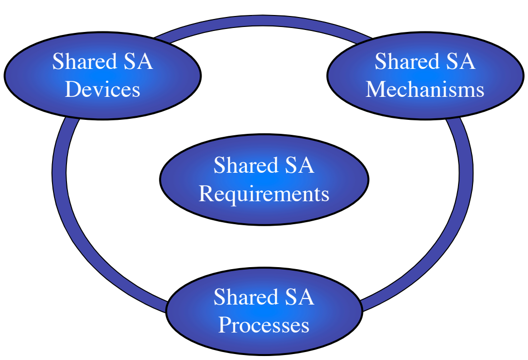


Figure 4. Factors Affecting Shared SA

Each of these factors — requirements, devices, mechanisms and processes — act to help build shared SA. Any measure or model of situation awareness also needs to take into account the effects of these factors on SA formation. These factors point out the need to include methods of SA information exchange in that it does not occur through verbal communication in our model. For example, proximity to other team members could allow for undetectable verbal glances and hand signals to occur between team members that would not occur with distributed teams.

1.3 Model of SA

Situation awareness is a dynamic construct that is based on multiple components. We have identified three main components that affect SA formation: an individual's abilities, the environment they are working in, and interactions with other team members (see Figure 5). Within each of these components are multiple factors that affect SA formation and maintenance.

Environmental factors take into account the physical environment of the teams. Factors such as distance between team members, size of the team, noise, stress and boredom can affect both an individual's and a team's ability to function at a high level. In addition, the interactions among the team members play a role in a team's SA. Teams that are working in the same location are more than likely to have higher shared SA then teams working distributively. For example, being in the same location allows for non-verbal cues to be exchanged amongst team members.

Most importantly, at the core of SA is the individual. His cognitive capabilities, skills and experience form the base for his SA and affects his ability to share the needed information with his fellow team members.

In order to measure SA we must first understand how these factors and processes affect the establishment and maintenance of SA in military teams. Figure 6 is a simplified model of SA formation. Individuals derive SA through various sources, as identified by the individual factors. A second major source of SA is information shared between individuals and their interaction with one another in an organization. A third source of factors that affect SA is the environment. Some of these factors are natural environmental features (e.g., location, proximity) or the soldier's personal condition (e.g., fatigue, boredom).

Each factor can seriously challenge the ability of the soldier to develop and maintain a high level of SA, and each can affect decision-making and action performance. We emphasize that SA comprises an iterative and dynamic process, as indicated by the arrows in the model. Individuals will make decisions and take actions based on their SA. Those actions will in turn affect the state of the environment itself (along with the action of other team members).

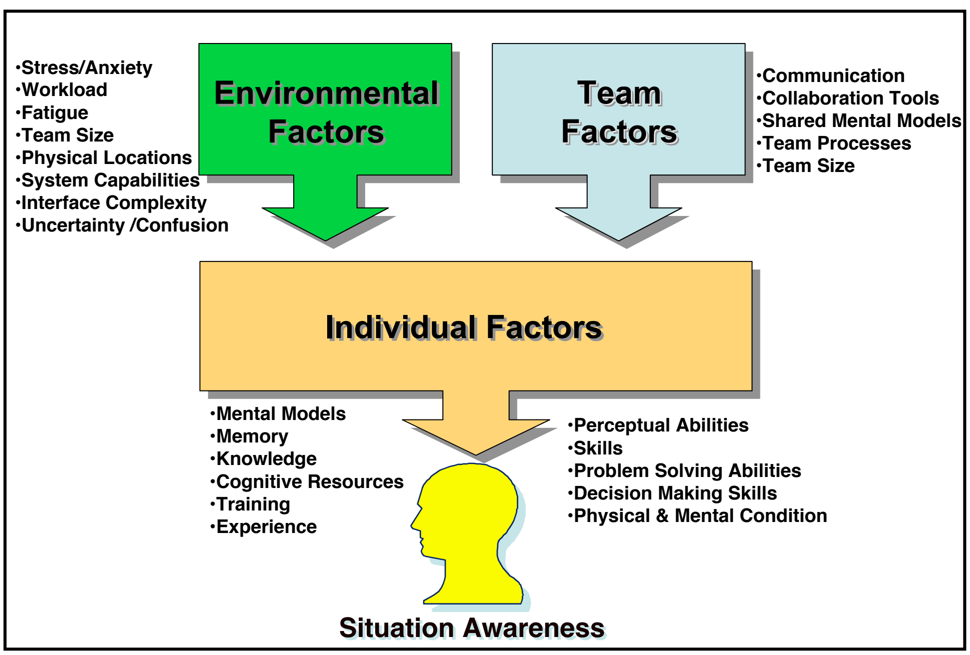


Figure 5: Factors Affecting SA Formation

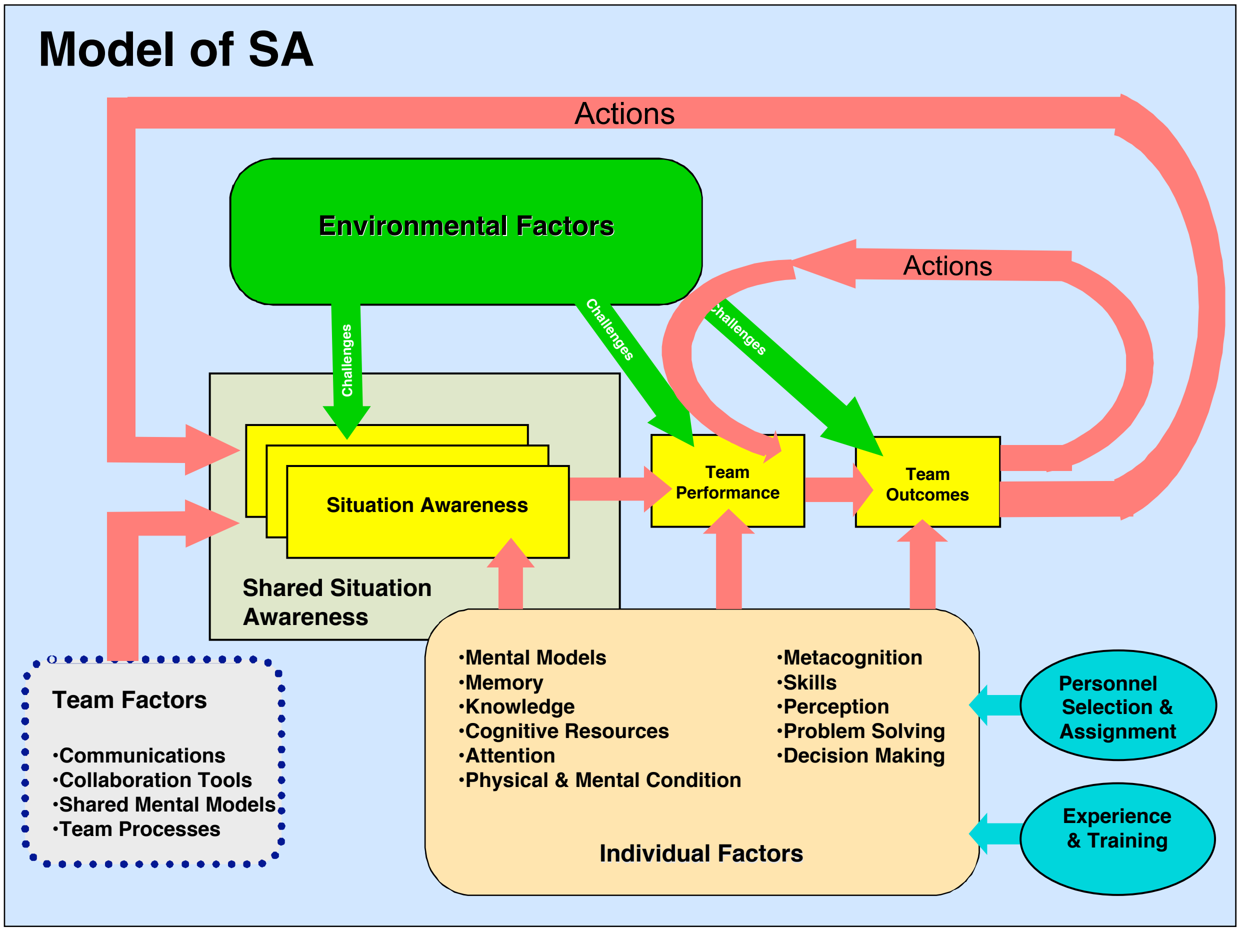


Figure 6. SA is a Dynamic and Complex Construct

Figure 6 shows that Situation Awareness is not a simple concept that can be measured be a single variable. Instead, any measure of SA must take into account the many factors that influence its formation. When measuring team and shared SA, the method and tools used for communication need to be addressed as well as the team processes and mental models used to share information between team members. As shown in the chart, in addition to the factors affecting SA formation so does outcome and actions of the individual and team. As indicated in the chart by the feedback arrows a negative outcome could be detrimental to a person's SA if it causes them to focus on the incorrect information.

1.4 Research Task Overview

Several tasks were completed in support of this research. The project activities accomplished during Phase I are identified in Table 1: (1) assess the potential teams and tasks (identify data collection and model validation needs); (2) assess and create the measures used in the model creation (review the modeling literature and select factors that are related to SA, mental models and shared cognition); (3) collect and analyze the data; (4) perform the Social Network Analysis (SNA) and create the shared SA model and (5) identify other factors that need to be included in the model to improve the SA prediction capabilities as well as identify model validation needs.

Our Team Met All Phase I Objectives

- Identify an unobtrusive SA measurement technique based on communication between individuals in a team.
- Select a measure that is unobtrusive, real-time and diagnostic.
- Validate this measure as a means of measuring shared SA.

Each of the Phase I task schedules is displayed below in Table 1. SA Technologies and Carnegie Mellon University completed all four major tasks of the project (Task 5 - model extension has not been approved for funding at this time), including identification of relevant data gathering exercises, identifying the factors that affect shared SA, shared mental models and shared cognition, creating the modeling databases, creating several initial models of shared SA and validating this model against SAGAT (situation awareness global assessment technique). Our modeling was based on current cognitive theory, social network analysis and state of the art SA measures.

A review of Table 1 shows that a special emphasis was placed on creating the necessary databases for further modeling in preparation for Phase II. As the data needed for this effort did not exist, we had to spend a considerable amount of time collecting this data.

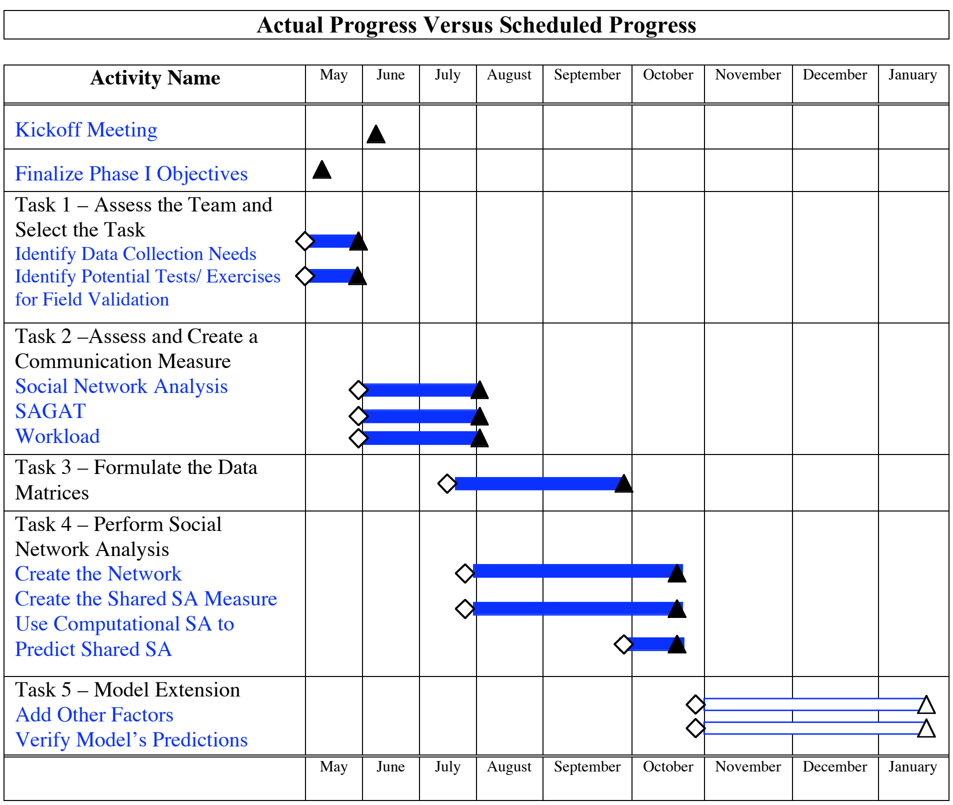


Table 1. Project Activities During Phase I

2.0 Task 1 – Assess the Potential Teams and Tasks

The objective of Task 1 was to assess the potential military teams and to decide where to collect the data needed to support our modeling efforts

While the goal is to produce a computational model of SA and ultimately a measure of SA that is global in nature, it is important that the initial scope of the project reflect current military operations and culture. We needed an organization that was large enough to have different teams that function in a distributed, nodal fashion as this is being introduced into modern operations. We also needed an organization in which shared SA and shared mental models are critical for good team performance. Lastly, we needed an organization that would be open to the data collection needs of our research. We were fortunate to have data sets available to use from two separate Army exercises that could be used for analysis purposes. In addition, we had a data collection opportunity at a joint service military command and control exercise, which allowed us the ability to collect the needed SA data.

Multiple Organizations for Data Collection Yields a More Robust Model

- Using teams of varying sizes and domains yields a model that is more applicable to other settings.
- Modeling large teams produces more stable models and enhances predictive validity.
- The model can easily be expanded to address individual SA, team SA and shared cognition.

2.1 Introduction

During the time of this research, we had the opportunity to analyze data collected from two separate military exercises sponsored by the U.S. Army. Both SA Technologies and Carnegie Mellon University collected the data through contracts with the Army Research Laboratory. While the main thrust of these exercises was not to support our modeling efforts, we were able to utilize most of the data collected on team communication and shared mental models and workload to aid in creating our model of SA. We also needed SA data in addition to the communication and shared mental model data. We collected the necessary data over the summer at the Joint Personnel Recovery Agency during two of their experiments. Since all of these events were conducted with very different organizations that had very different goals, the data collection methods used across the events are slightly different.

2.2 Important Findings

- A review of databases collected from other team studies did not yield the necessary information needed for our models. Social Network Analysis is a relatively new concept and thus current research databases do not have the parameters used in this domain. For this reason, we had to use data collected from recent military exercises as well as collect data on SA and communication at a military command and control exercise.

- There has been little investigation of team communication behavior in C^3I organizations, such as communication patterns between individuals, dynamical forming teams based on communication, and reciprocity of communication among team members.

- One cannot model SA without first having the data needed to create the parameters that go into the model. We believe that a single factor model will not yield good predictions and thus had to find a method of measuring SA that would allow us to utilize multiple types of data.

2.3 Methodology

We put together a list of criteria for selecting our databases and venues, to ensure the data we collected would fit the needs of our model and could ultimately be used to validate the model in terms of its ability to measure shared SA.

- Nodal organization
- Dynamic teams
- Moderate to large size
- Shared SA plays a role in team performance
- Willingness to collect multiple data types during stops in the exercise (for actual collection events)

2.4 Results

As a result of our efforts we were able to secure datasets from two different Army venues as well as arrange for data collection to occur at a third joint service event. Each of these data collection events will be described in more detail.

<u>United States Army Future Force.</u>
The Army is undergoing a transformation from a traditional hierarchical structure to a more nodal distributed organization called the Future Force. In order to test the effectiveness of this organization and to determine if the new assigned roles and tasking are appropriate, several exercises have been conducted at the Army's Battle Labs. We analyzed data from two of these exercises: a smaller exercise of 56 soldiers from the Fort Leavenworth Battle Command Battle Lab and much larger exercise from the Unit of Action Maneuver Battle Lab at Fort Knox, which consisted of over 250 soldiers in 6 different locations.

<u>Joint Personnel Recovery Agency</u>
The Joint Personnel Recovery Agency (JPRA) is a subordinate activity of U.S. Joint Forces Command. As the Department of Defense (DoD) executive agent for personnel recovery, JPRA is responsible for the shaping, planning, preparation, execution and repatriation of personnel recovery. Recover centers are staffed all over the world. For this exercise, data was collected at the Personnel Recovery Education and Training Center, where they train servicemen to staff the recovery centers. The servicemen, who are composed of Navy, Army, Marines and Air Force personnel, attend a two week program in which they receive one week of course work followed by a one week simulated exercise. During this exercise each servicemen gets a chance to work at one of the four different recovery centers: Navy, Army, Special Operations, and a Joint Search and Rescue. The exercise is designed to mimic real life events in a recovery center. Data was collected from two separate exercises conducted in June and July 2004.

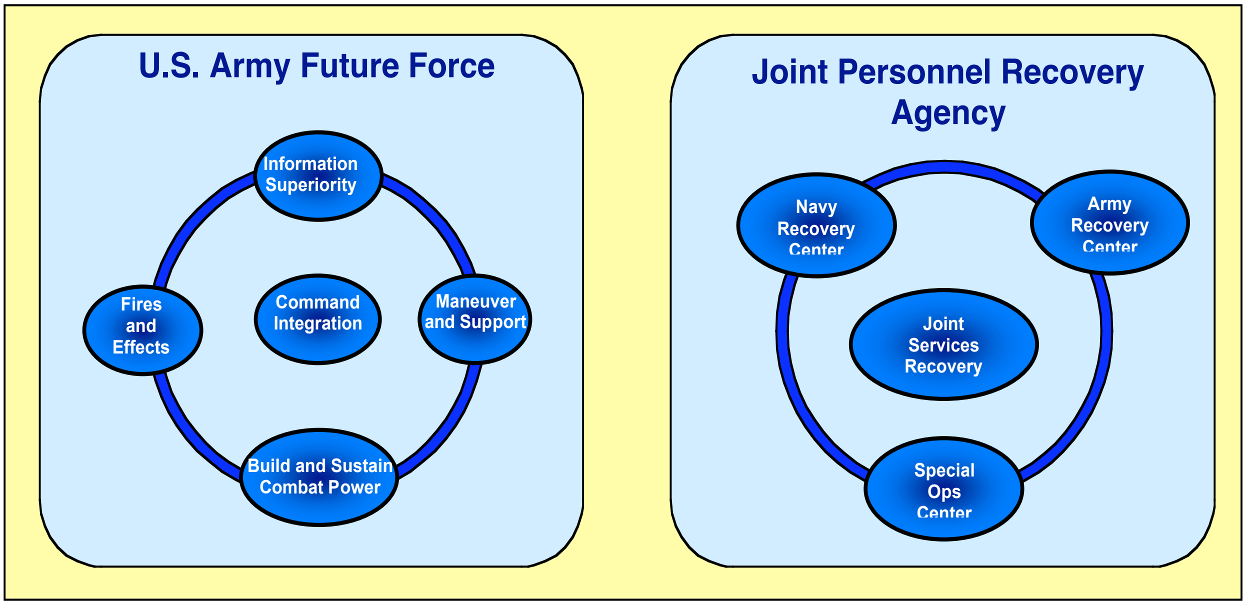


Figure 7. Data Collection Locations

Figure 7 illustrates the structures of the organizations in which data was analyzed for our modeling efforts. Both organizations use a nodal teamwork structure as opposed to a hierarchical organization. The teams in each organization vary in size, but the need to communicate and share information is identical. Both models and measures of shared SA can be very beneficial to both of these organizations.

3.0 Task 2 – Assess and Create a Communication Measure

Using Social Network Analysis and Modeling Leads to Better SA Measures

- Social Network Analysis (SNA) provides a graphical view of how SA is formed in the organization.
- Modeling allows for multiple factors to be included leading to a more accurate SA predictor.
- Once an acceptable SA model is created, the model factors can be measured, real-time, for an unobtrusive measure of SA.

During this task we both created the data collection methods for the individual exercise and analyzed the existing data. In order to support our modeling efforts, three different kinds of data were needed.

- SA Measures – SAGAT Data
- Shared Mental Model Measures
- Social Network Measures

As mentioned earlier the methods were modified slightly to fit each of the experimental domains and size. Each of the methods will be described in more detail below.

3.1 Introduction

The modeling efforts for this research are based on the domain of Social Network Analysis (SNA). Social Network Analysis (SNA) is a method designed to focus analysis on a network-based view of the relationships between people and organizations (Dekker, 2002). SNA allows for the quantification of dyadic ties that exist among team members. In any organization or team people influence each other, the ideas being exchanged and the flow of information (Borgatti, 2002). Thus, a social network is not just a description of who is in the team but how they are put together and how they interact with one another (Borgatti, 2002). In addition, SNA allows for values to be attached to these relationships to represent strength of the relationships, information capacity, rates or flow of traffic, distance between nodes, probabilities of information being passed (Borgatti, 2002). It is these values that allow SNA to quantify the relationships thus presenting a means for mathematically testing the network.

SNA is performed using a measure of communication frequency collected on every individual in an organization. Other measures can be included in the resulting model such as workload, experience and other factors deemed as potential predictors of the variables of interest. In our case, we are interested in using Social Network Analysis, Computational Modeling, and graph theory to assess shared SA and mental model congruency, and to also use the data to simulate how the organization would behave (without actually running role-players). This approach we are using has the advantage of helping us develop the ability to not only measure SA, but to also project the shared SA of the organizational members. Trying to determine what factors should be included in this model in addition to our communication measure, was part of our phase I efforts.

3.2 Important Findings

- Creating a model or measure of SA needs to include more than just communication. According to Endsley and Jones (1997), other factors such as team devices (how the information was conveyed e.g. verbal or non-verbal), and team processes (e.g. did they read back or cross check with one another) affects the formation of shared and team SA.

- While effective team communication is essential to team performance the quantity of information is not a direct indicator of performance (Fjelde & Switzer, 1994). Mosier and Chidester (1991) found that better performing aircrews actually communicated less than poorer performing ones.

- In order to validate our model and the factors selected as measures of SA, we collected objective SA during a C^{3}I exercise. The measure we selected was SAGAT (situational awareness global assessment technique). SAGAT is one of the only objective measures of SA that has been found to be both valid and reliable (Endsley, 1995a).

- Because of the sheer complexity and size of the organizations we selected for data collection and analysis, we were unable to gather measures of the same data types at each event.

3.3 Methodology

Data was analyzed from two separate military exercises: Fort Leavenworth and Fort Knox. Data from both of these experiments was collected for other research purposes, but we were able to use some of the data collected for our model. Another experiment was conducted at the Joint Personnel Recovery Agency to gather the needed SA data. In addition, at this event, we also collected communication data and social network data. In all three exercises, background information such as rank, years of experience and current job were recorded. This is in addition to the data types listed below.

3.3.1 Methodology Ft. Leavenworth

The Fort Leavenworth Battle Command Battle Laboratory (BCBL) gathered fifty-six army officers to serve as role-players for an experimental command and control staff (Figure 8). Each role-player was assigned to a functional cell with three to eight other role-players. The role-players gathered information, coordinated with appropriate staff members, and entered battlefield actions into the simulation. Partitions or walls separated the seven cells, so that a participant could talk directly to members of his own cell, but could only communicate with members of other cells using the communication tools provided to them.

Data Collection occurred constantly and in multiple forms throughout this exercise. Critical to the analysis was an automated self-report collection system that was executed every 60-90 minutes during the simulation. Data was collected using a networked questionnaire that asked the participants for feedback regarding the prior session. Questionnaire data was collected for a total of 16 sessions. For the 7-11 minutes that the data was collected, the simulation was frozen until all responses were recorded.

Figure 8: Command Post Exercise at Fort Leavenworth, Kansas.

<u>Social Network Data</u>

Social network data was gathered by asking the participants to report the people they had communicated with in the time since the previous questionnaire. They could give up to 10 responses by selecting participants from pull-down menus. The responses were ordered by the frequency of communication during the previous session. They were asked to give a rating of 1 to the person they talked to the most and 2 to the person they talked to the second most up to the 10th most frequent person they talked to.

Shared Mental Model Estimates
Shared mental model estimates were gathered using the NASA TLX (Task Load Index) assessment consisting of six workload parameters on a Likert scale (Hart & Staveland, 1988). The parameters are mental demand, temporal demand, effort, own performance, frustration level, and physical demand (see Figure 9). Participants were asked to rate themselves as well as five other people randomly selected from the other participants. When rating other people, participants had the option of selecting "Don't Know" for each of the six questions.

We used Entin and Entin's (2001) proximate measure of shared mental models in this study. They found congruence between participants' mental models and their ability to rate other teammates workload.

Mental model congruence was determined by comparing each person's self-reported workload with the estimation of that person's workload by other participants. This measure was computed by summing the absolute differences between the self-reported ratings and the rater's estimations. For example, if person A's self report was a 5 for each question on the index, and person B estimated A's workload as a 3 for each question, person B's mental model congruence would be 12 (two multiplied by six). Congruence scores could range from 0 (indicating perfect congruence) to 36.

Figure 9: Workload Questionnaire Administered During the Fort Leavenworth Exercise

3.3.2 Methodology Ft. Knox

In June 2004 the U.S. Army began a one-month simulation exercise to study the effectiveness of a new method of organizing Army staff personnel known as the Unit of Action (UA). The Unit of Action is intended to replace the traditional Battalion organization with a more flexible design capable of adapting to dynamically evolving situations. About 250 active duty and retired soldiers participated in the exercise at 6 locations distributed throughout the United States. The participants could communicate with their remote colleagues via email or radio network. During the exercise, participants completed an on-line survey. All answers were based on the time period since the last survey was collected. The survey was implemented as a web form, which the participants completed in an ordinary web browser. All answers were multiple choice.

Social Network Data
During the exercise, participants were asked to list the top 7 people they communicated with (in descending order). This is the same method used in the Ft. Leavenworth exercise except only 7 people are selected as opposed to 10. The communication survey was filled out by all participants 2-4 times per day, depending upon the pace of the operation.

Using the communication data, we constructed a social network graph for every session. The social network graphs were used to calculate the social network distance (geodesic) between each player.

Workload Ratings
During this exercise participants only completed workload rating for themselves using the NASA-TLX.

SA Congruency
Due to the size of the exercise, we were unable to administer SAGAT queries to all the participants. We therefore elected to create an SA congruency measure which is more reasonable for such a large organization. Our SA congruency measure is based on Endsley's (1995b) level 3 SA or the projection of what will happen in the environment. Participants were asked "What are the three most likely risks to this operation in the immediate future" at each stop. Using this data, we constructed a pair-wise measure for each pair of participant responses, which estimated the similarity of their risks. This provided us a measure of risk congruency of similarity.

3.3.3 Methodology Joint Personnel Recovery Agency

In June and July of 2004 we gathered data from two separate exercises at the Joint Personnel Recovery Agency (JPRA). All four service branches were represented in these exercises. This is the only exercise that we actually attended to collected data for this SBIR research. The other events occurred earlier and we only analyzed the existing data.

Each of these two exercises consisted of 17 players and 5 different scenarios over a three-day period. During the simulated exercises the scenarios were stopped three times to collect data for a total of 15 stops.

Figure 10: The Joint Personnel Recovery Agency

Situation Awareness Measure
During this exercises the Situation Awareness Global Assessment Technique (SAGAT) was used. It is an objective measure of situation awareness (SA) (Endsley, 1995a). SAGAT is designed to elicit information from all three levels of SA – perception, comprehension, and projection. At random times, the exercise is stopped and SAGAT is administered to all the participants.

The foundation of a successful SAGAT data collection effort rests on the efficacy of the queries. Before queries can be developed, the operators' SA requirements must be defined. This task is accomplished through a goal-directed task analysis (GDTA). The GDTA seeks to uncover the goals operators have in a particular domain, the decisions that must be made to achieve these goals, and the dynamic information requirements needed to support the decisions. (For more information on GDTA, see (Endsley, 1993). The SAGAT queries are based on these information requirements.

In June, six instructors were interviewed at the JPRA. These interviews were one on one and lasted approximately two hours. The interview notes were turned into a GDTA that was used for the construction of the SAGAT queries for the upcoming exercises. Based on the fidelity of the simulator and the criticality of certain information requirements as identified by the instructors, seven queries were created for these exercises. The queries are shown in Appendix A.

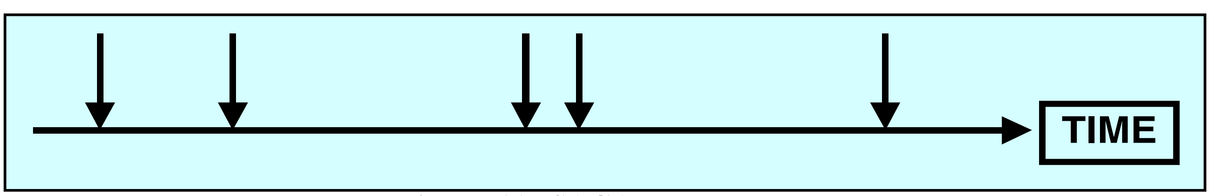


Figure 11: SAGAT Methodology

The SAGAT methodology involves stopping the exercise/simulation at random points in time and administering a rapid battery of queries to ascertain the subject's SA at that point in time and then scoring the subject's SA based on the objective data obtained from the exercise/simulation.

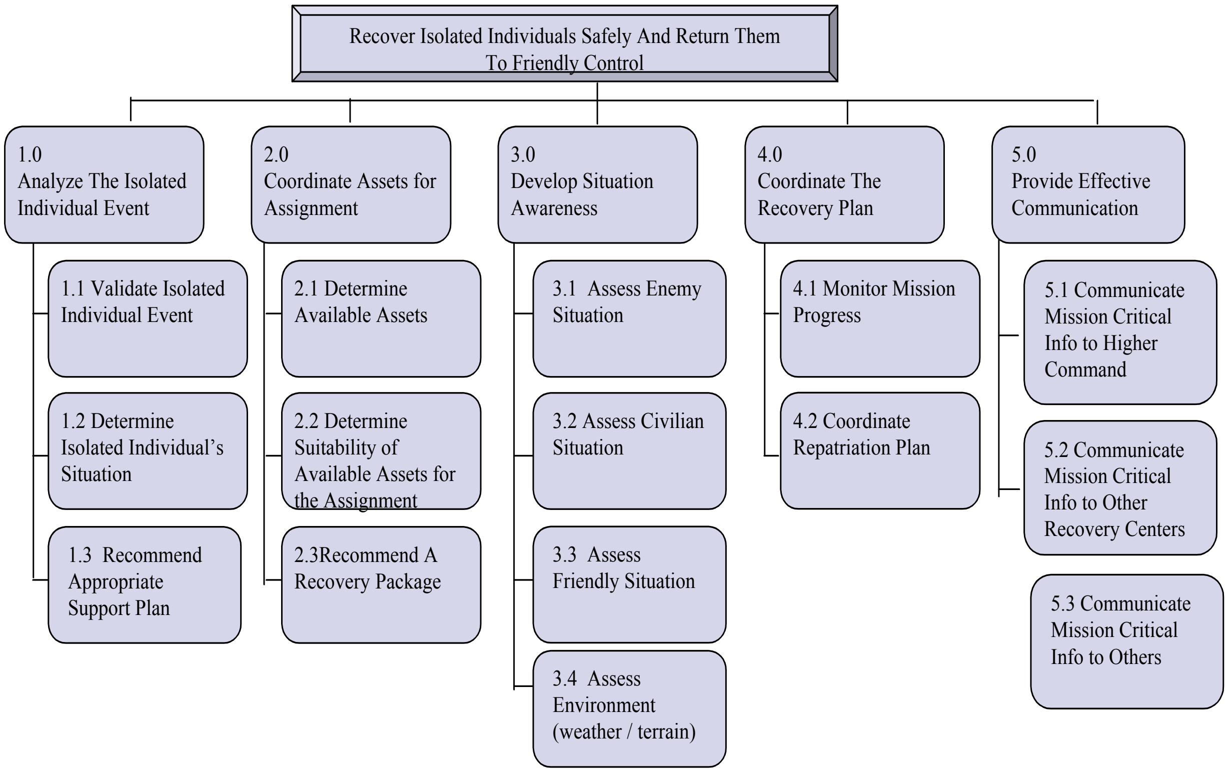


Figure 12: Goal Directed Cognitive Task Analysis for the Personnel Recovery Center

Figure 12 shows the top-level goal for members of a personnel recovery center. In a GDTA the major goals of a particular position are identified, along with subgoals necessary for meeting each goal. Associated with each subgoal, the major decisions to be made are identified. The SA needs for making these decisions and carrying out each subgoal are subsequently identified.

Shared Mental Model Estimates

Shared mental model estimates were gathered using the NASA TLX (Task Load Index) assessment consisting of six workload parameters on a Likert scale (Hart & Staveland, 1988). The format was similar to the Fort Leavenworth exercise except participants were asked to rate themselves and each of the other 16 participants at each stop. The measure was administered using paper and pencil. Mental model congruence was determined in the same manner described previously.

Social Network Data

During the exercise, participants were asked to rate the top 4 people they communicated with (in descending order). All participants were listed on the social network data sheets including the white cell players. During this exercise, the white cell players also completed a network sheet and they were asked to rank the top four individuals they communicated with since the last stop.

4.0 Task 3 – Formulate the Data Matrices

Data Matrices Form The Basis For The Modeling Efforts

- The matrices used in the modeling efforts were created from the data collected during the three different exercises.
- The matrices form the foundation or databases used in modeling.
- The creation of the matrices was an iterative process as more factors were added to our SA models.

The critical first step in a social network-based approach is to construct the network itself. Command and control environments, however, provide an interesting hurdle to traditional social network based approaches. In traditional social network approach, the organization is relatively static and predominantly shaped by its formal organizational chart. In a command and control environment, the network is constantly shifting and reorganizing and the formal organizational chart is quickly discarded for a dynamic, task-based structure. Driven by the shifting priorities and tasks, a command and control organization becomes a network of forming and disbanding ad hoc teams and shifting leadership (Graham, Gonzalez & Schneider, under review).

We chose to expand social network theory into a concept we have labeled Dynamic network analysis. Dynamic network analysis considers the organization form to be a living entity capable of shifting form and structure (Carley, 2003). Dynamic network analysis also encompasses a methodology to mathematically represent an organization and its member connections through linked matrices (Carley, Ren, & Krackhardt, 2000). These matrices can include things such as communication frequencies, individual experience, information needs, current tasking, level and type of information requirements exchanged.

4.1 Introduction

Dynamic network analysis data collection is reliant upon communication data. Communication data can be gathered by shared email headers, chat room traffic, instant messaging, phone calls, or by surveying the individuals (Wasserman & Faust, 1994). While each of these communication mediums has different qualities, measure relevance is determined by the organizational context and collaborative tool characteristics. For instance, in one study, we found that early in an organizations life, people are more comfortable with face-to-face, but as they become comfortable with collaborative tools, they migrate their important communications to those tools. As researchers, we would need to track both communication mediums until a full transition to one or the other occurs. In the studies specifically referenced in this chapter, we describe both chat and self report data.

While we would have preferred to log all communications regardless of medium, the computer systems for our various experiments (Ft. Leavenworth, Ft. Knox, JPRC) would not support an automated logger. Plus, the time constraints of this phase I SBIR would not allow us to analyze such data. As a result, a self-report questionnaire was employed as a simplified method for collecting all potential communications between members regardless of the communication method used. The self reported questionnaire was validated last year during an Army exercise in which we could collect both self-report and chat data. Therefore, we utilized the same methodology in our Phase I research described in this report.

4.2 Important Findings From Our Earlier Validation Work

- Translating real-time collaboration data into network graphs and data matrices is achievable. However, the translation software is specific to the construction of the collaboration software in use. (ie. A translator for DCTS raw data will not support MCS raw data without minor modification ~3-7 hours of programming time)

- Using only one ‘channel’ of a collaboration log (chat, email, etc) will not provide an accurate representation of individual communications networks. Our earlier work indicated that as much as 65% of the communication is missed by only monitoring one type of communication (see Appendix B). However, using only one channel of a collaboration log will provide lesser scale, but similar pattern group/organizational communications networks.

- A network-based approach to individual situation awareness must consider all communication channels and will have limited to no validity using only one communications channel. However, a network-based approach to group or shared situation awareness dependent upon one communication channel will have validity in modeling SSA trend and limited validity in modeling shared SA amount (see Appendix B for a description of this earlier validation work).

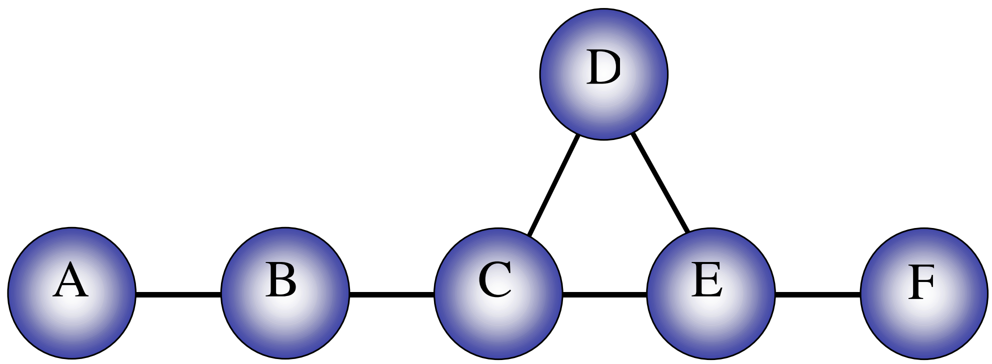


Figure 13: Simple Social Network Graph

A social network is a graph consisting of individuals and connections between them. In a social network graph, individuals are represented as nodes and communication between individuals is represented by links between the nodes (Borgatti, 1994). Social network distance (often referred to as a geodesic) is the number of links or actors between two members of a social network graph.

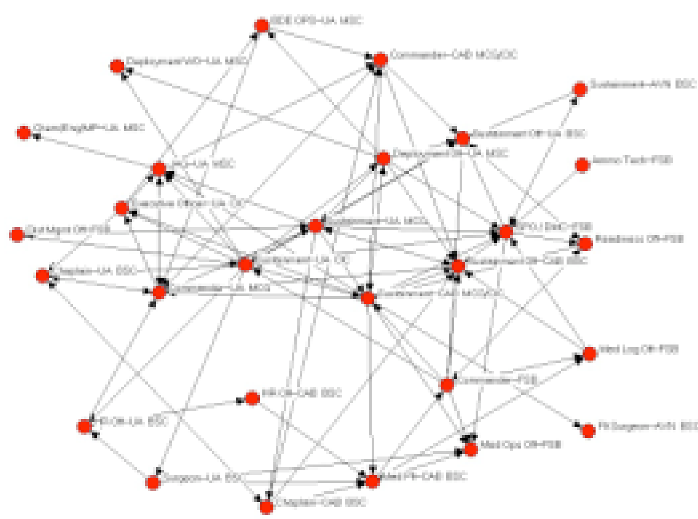


Figure 14: Complex Social Network Graph

This graph shows data collected from twenty-eight experienced army officers at an Army command and control exercise.

5.0 Task 4 – Perform Social Network Analysis

> **In Building Our Model An Iterative Approach Worked Best**
>
> - Our models were built iteratively adding factors from the SA model that have a large impact on SA formation.
> - The simplest model: shared SA = direct communication was very weak.
> - Adding in other factors produced a much more robust model with a very high prediction rate (r^2).

5.1 Introduction

Our approach rests on the belief that SA is not a simple construct that can be measured by only one predictive variable, such as communication. This is obvious in the model of SA presented earlier. Rather multiple factors need to be considered when measuring SA, which is why our approach is complex. The complexity comes from considering all of the contributing factors to a single person's situation awareness and factors that contribute to any two organizational members' shared SA. Factors such as geographical distance, leadership, collaborative tools, network proximity, acknowledgement, familiarity and others all need to be considered for inclusion in any measure or model of SA. Obviously, this is not possible in a Phase 1 research project and we have only begun to address which of the factors has the greatest influence on SA formation. In this section we present our initial models that look at four main factors: direct communication (who communicated with who), physical proximity (distance between individuals), network leadership (who was communicated to most often), and homophilly (similarity of backgrounds).

For this research, we focused on shared SA amongst the team members. Shared SA is a reflection of how similar the team players view the current situation. Thus, if a team has a high degree of shared SA we can assume they are interpreting the information requirements in a similar manner. In future research we will address team SA, which is how well each member knows the information requirements needed to perform their own tasks. We felt that shared SA provides the clearest indication of a team's functioning together and therefore focused our phase 1 efforts here.

5.2 Methodology

We took an iterative approach to developing a shared situation awareness metric. In doing so, we deconstructed shared situation awareness into the situation model and in particular shared situation awareness level III. First we established a benchmark metric to evaluate each iterative metric's performance. Next, we developed a metric to understand how far out from each participant we could expect their situation model to extend and validated that model against the data collected from the experiment at Ft. Leavenworth. Last, we applied the situation model metric to the shared situation awareness metric (level III) development and validated this model against the data collected at the Ft. Knox experiment.

5.3 Technical Strategy

Our technical strategy is to use an iterative approach in which we start with the simplest of models and add more factors to increase the accuracy of the model in terms of predicting shared SA. We begin with a simple model that states that SA is directly measured by communication, next we test a model with 3 factors and then we test a model with 4 factors. Each of these models is described below. See Appendix C for more detailed information on our modeling efforts. For these initial modeling efforts we used a measure of shared cognition developed by Entin (2001) instead of shared SA due to the limitations of the databases. Ultimately, however, the final model was validated against an SA measure.

5.3.1 Baseline Model (Model 1)

Methodology

As a baseline for model efforts we used a simple model that states that SA is directly measured by communication. The only variable used in this model is direct communication. In essence this is self-reported measure of who reported talking to whom. Direct communication between organizational members has been repeatedly demonstrated as a key variable in the development of shared mental models, shared situation awareness, and transactive memory. Many approaches to shared mental model and shared situation awareness rely solely upon direct communication to estimate the 'sharedness' between two organizational members. In terms of the algorithm, shared situation awareness (SSA) between two organizational members (i,j) is a function of whether or not they have directly communicated (Dij) during the time period of interest. Our goal in using this as a benchmark is to progressively beat the results achieved with a more informed Shared Situation Awareness model.

$$SSA_{ij} = \square D_{ij}$$

Validation

We validated this model using both the Ft. Leavenworth data and the Ft. Knox data. In our first test of the model using just our communication survey the model was only able to account for (9 % of the variance (F(1,19933) = 4.24, p=.039). In our second test we included all of the communication's that occur between two individuals on all communication channels (chat, face to face, email and voice) and we still had a poor estimate of shared cognition. Using this model we only achieved (at best) a 15% accuracy rate, F(1,631)= 5.58, p=.018

Model 1	$D_{ij} : DirectCommunication$
Model 2	$P_{ij} : Physical\Pr oximity$ $C_{ij} : Geodesic(SocialNetworkDist)$ $A_{ij} : NetworkLeadership$
Model 3	$P_{ij} : Physical\Pr oximity$ $C_{ij} : Geodesic(SocialNetworkDist)$ $A_{ij} : NetworkLeadership$ $H_{ij} : Homophilly$

Figure 15: Factors Included In The SA Models

5.3.2 Model 2

Since the ultimate goal of our research is to develop a model of shared situation awareness between members of an organization, the situation model we wanted to understand is that of the organization itself. The model we selected incorporates the value of different system parameters and includes an understanding of the dynamics of the system.

We developed a hypothesized model after an extensive review of the literature. Our original literature review and experience indicated that an individual's situation model of an organization is a function of physical proximity, network distance (nodes on the geodesic) & organization communication status (authoritativeness). These three factors were included in this model. In this instance, we were using a person's situation model of the organization instead of shared SA. If this model is selected as our best fit, we will validate this model against actual SA data.

$$SituationModel_{ij} = \square A_{ij} + \square P_{ij} + \square C_{ij}$$

5.3.2.1 Model 2 Factors

Physical Proximity

Physical proximity has been found to favor the development of models of others and improves performance. Through this observation, team members are able to more accurately obtain information about other's capabilities, tasks, and situation and are better able to establish and maintain a situation model of the people they interact with. In the case of Graham, Schneider, and Gonzalez (2004), we found that physical collocation was twice as likely to produce a shared mental model. We measure physical distance based on the metric distance between individuals (i,j) in the organization. If two members are physically collocated, we consider this a distance of zero. As they become more geographically dispersed, so does physical distance.

Geodasic (Social Network Distance)

Multiple studies have also found that communication supports situation model development (Salas, Rozell, Driskell, & Mullen, 1999). Team members that communicate directly communicate tend to understand each others tasks and situation and are able to gather information about the other's capabilities (Graham et al., 2004). We extend the definition of communications beyond direct communications to include the chain of communication in terms of the number of nodes on the geodesic between two agents. We measure network distance based on the number of edges in the geodesic between two members of the organization. The geodesic is the shortest number of edges between two members (i, j). An edge is a communication link between two members of the organization. Even if two organizational members do not directly communicate, there is a likely set of communication links with other members that will connect them.

Network Leadership

Members in close proximity to a leader are in the military C2 culture, more likely to have a good situation model of their leader than other organizational members at an equal distance. This phenomenon occurs because, in the military C2 culture, leaders are expected to have the most correct situation awareness (French & Hutchinson, 2002) and explicitly state their assessment of the situation and provide their intent for future activities to their immediate leaders and subordinates. Network-based informal leadership is measured through the eigenvector centrality in the dynamic communications network (Scott, 1992). This descriptor of leadership assigns members with higher eigenvector centrality as leaders of leaders, and members lowest in eigenvector centrality hold strictly subordinate roles.

5.3.2.2 Model 2 Validation

Figure 16 is a graph of the situation model metric accuracy over the course of the experiment as compared to the mean performance of the simple direct communication measure (Dij) as a baseline for performance. The mean baseline performance (Dij) only accounted for 15% of the variance in field measured situation model accuracy during the experiment (see the yellow line). The mean r-square for the situation model accuracy metric was .24 ($p<.001$; $F(3, 2298) = 564$). The metric performance steadily improved as applied to organizational data collected later in the experiment with its best performance accounting for 41% of the variance. The situation model metric clearly outperformed the baseline for metric performance.

Our situation model metric performed well. Any time a researcher finds a metric that accounts for 30-40% of the variance of any variable in a large organization is considered a publishable result. Further, we nearly doubled the performance over the baseline metric of direct communication. However, for a military real-time application, we need performance to be in the 70-80% accuracy range. The third iteration of our metric adjusted for the lessons learned and observations from the first metric iteration.

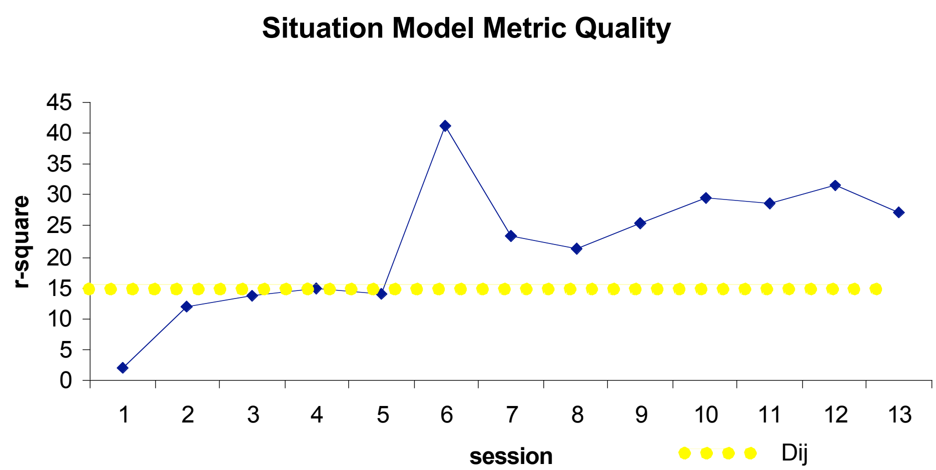


Figure 16: Comparison of the Baseline Model Against Model 2

Figure 16 is a graph of the quality of the situation model metric over the course of the experiment. Dij is the mean performance of the direct communication baseline metric for comparison.

5.3.3 Model 3

Model number 3 was validated against a metric for shared situation awareness level 3. Additionally, the third model took advantage of an observation made during the post-hoc analysis of the second model. Specifically we found that the organizational member made significantly more accurate workload estimates of organizational members with similar backgrounds as themselves ($p<.01$; $F(29, 1539)= 22.96$). Background similarity, in this case, considers years of service, branch of training, and types of staff experience/assignments.

In the social network literature, background similarity has strong connections with the concept of homophilly. Homophilly theory states that members are more likely to create communication ties with other group members who they deem to be similar. In colloquial terms, "birds of a feather flock together." Brass (1995) observes that "similarity is thought to ease communication, increase predictability of behavior, and foster trust and reciprocity". Work by Espinosa, Slaughter, Herbsleb, Kraut, Lerch, and Mockus (2001) demonstrated that background familiarity improves the shared mental model between members of a team. In this case, we are not using homophilly to estimate the likelihood that two people will communicate, but instead we are seeking to estimate the shared situation awareness between two people in an organization. Hij represents a background similarity score between any two organizational members (i,j). Homophilly was calculated based on a similarity score from background information the participants provided in their user profile.

$$SharedSA_{ij} = \square A_{ij} + \square P_{ij} + \left(\frac{\square H_{ij}}{\square C_{ij}} \right)$$

5.3.3.1 Model 3 Validation

The shared situation awareness algorithm was validated against a data set collected at an organizational experiment conducted at Ft Knox, Kentucky. The data set is from a trial 256 member command and control organization. The role-players gathered information, coordinated with appropriate staff members, and entered battlefield recommendations/decisions. The participants could communicate with their remote colleagues via email or radio network. During the exercise, participants completed an on-line survey. All answers were based on the time period since the last survey was collected. The survey was implemented as a web form, which the participants completed in an ordinary web browser. To reduce interruptions during the scenarios, all answers were multiple choice.

<u>Shared SA (level III) Field Measure</u>

The best validation of our metric would be against a congruency in SAGAT scores between each participant. A SAGAT would require the participants to provide extensive information about their perceptions, comprehension, and projections relative to the current environment and situation. However, due to the size and pace of the exercise, we were unable to administer a full SAGAT at every collection period. We were, however, able to employ a SA congruency measure to account for Endsley's (1995b) level 3 SA or the projection of what will happen in the environment.

To find congruence in level 3 SA, participants were asked "What are the two top risks to this operation in the immediate future" at each stop. They could choose from a total of twenty-two choices that were divided into categories of Friendly, Enemy, and Environment. Using this data, we constructed a congruence score for all pairs of organizational members.

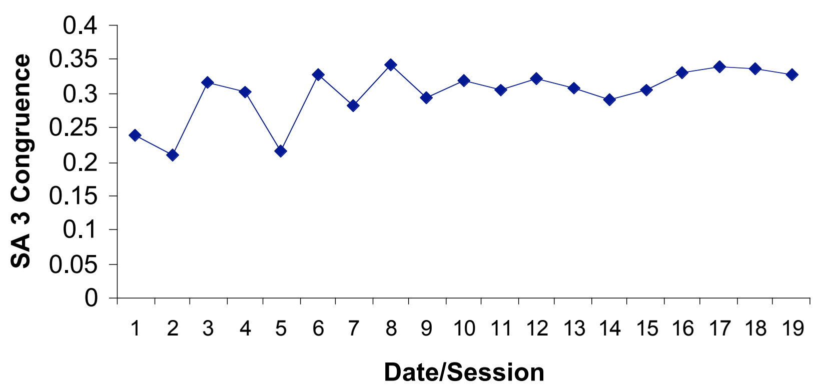


Figure 17. Model 3 - level 3 Congruence Over the Duration of the Experiment

Figure 17 is the mean situation awareness (level 3) congruence for the organization. In the early stages of the experiment, there were tremendous fluctuations in congruence as the organization trained and the individuals learned their roles.

Metric Validation

Figure 18 is a graph of the quality of the shared situation awareness metric accuracy over the course of the experiment as compared to the mean performance of the simple direct communication metric (Dij) as a baseline for performance. The mean baseline performance (Dij) only accounted for 9% in field measured Situation Awareness (level 3) congruence during the experiment. The mean r-square for the shared situation awareness metric was .78. The metric performance range fell between 58% and 98% over the course of the experiment. The shared situation awareness metric clearly outperformed the direct communication baseline.

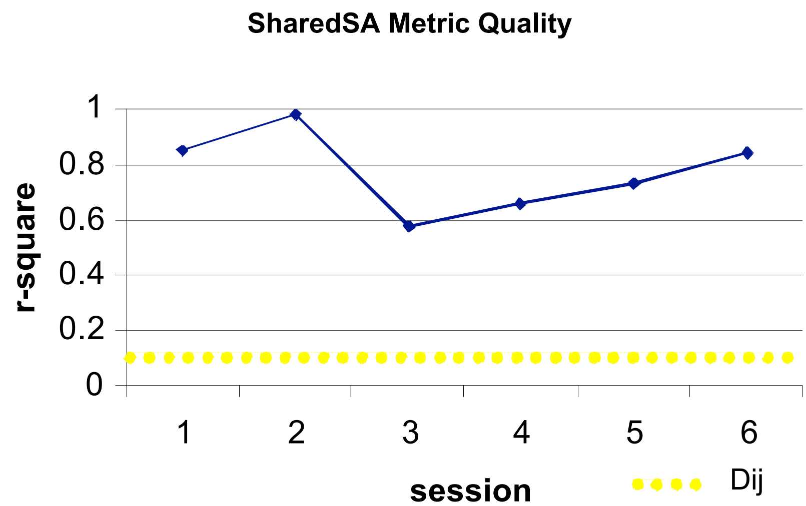


Figure 18: Validation Results of Model 3

5.4 Result

The results from our modeling work clearly indicate that the model with the shared situation awareness metric performed extremely well (model 3). Accounting for .78 of the variance in a model is considered very good. Figure 19 shows how much variance was accounted for by the factors included in the models.

One of the key things learned while performing this research was how to cut the analysis time down from weeks to minutes. Figure 20 shows our progress in this area.

We do however, realize that our model can be improved and should also be validated against level 1 and level 2 SA and team SA.

What We Created With Our Modeling Efforts

- Non invasive model of Shared SA
- Implementable now: can capitalize on basic C3I data structures.
- Validated on multiple large scale military C2 organizations

Model 1	$SharedSA_{ij} = \square D_{ij}$	$r^2 = .09 - .15$
Model 2	$SituationModel_{ij} = \square A_{ij} + \square P_{ij} + \square C_{ij}$	$r^2 = .24$
Model 3	$SharedSA_{ij} = \square A_{ij} + \square P_{ij} + \left[\frac{\square H_{ij}}{\square C_{ij}}\right]$	$r^2 = .78$

Figure 19. Validation Results of Our Models

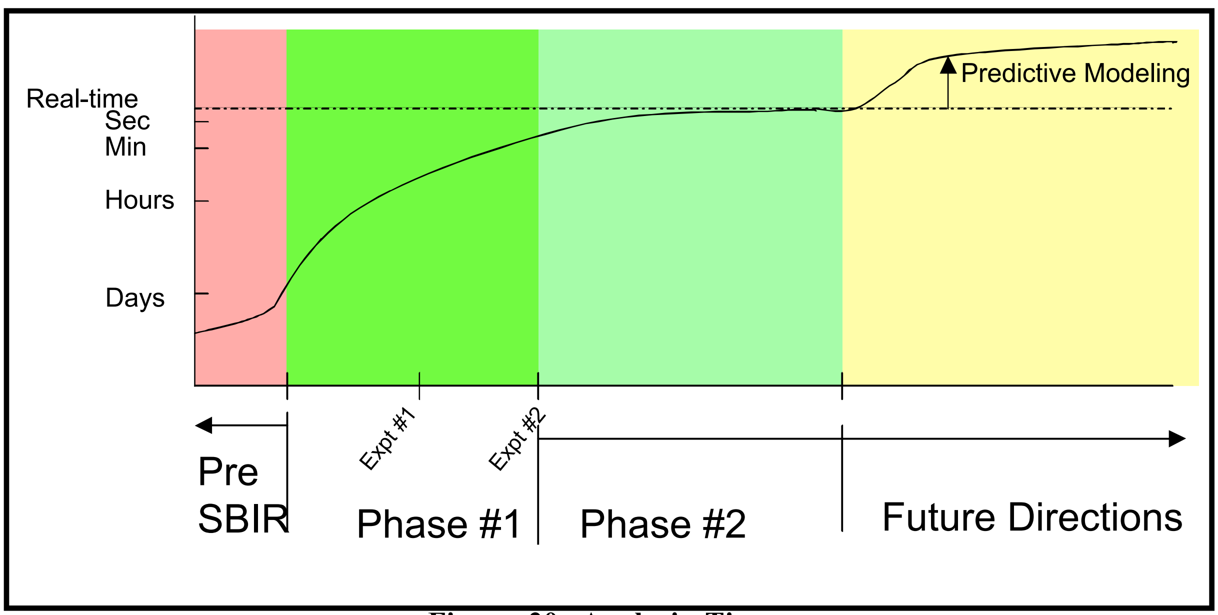


Figure 20. Analysis Time

Figure 20 shows how we cut the shared SA analysis time from days to hours to minutes after data log is collected.

6.0 Research Accomplishments

Our research strove to create both a model of shared SA as well as an unobtrusive measure of SA. By first understanding the various factors that make up shared SA we created a more robust predictive model of SA using the identified factors. The methods used to gather the data for model validation can also be used as real-time SA assessment measures (e.g. communication surveys, workload). Initially, it took us several days to output our shared SA measure. However, in our research we were able to turn the data around in under 20 minutes at the conclusion of this research project. While not real-time, we believe we can get SA measurement results in 20 seconds or less with just a few more months of research. Further, the metric we have developed is sufficiently valid for application to real-world shared SA tracking in military command and control organizations.

This section highlights the Phase I research accomplishment. Each accomplishment references report sections and/or appendices that describe and elaborate upon the accomplishment.

- Established several databases of SA, communication and workload measures. See 2.4 Results
- Identified key factors that influence SA formation. See 1.3 Model of SA
- Created a Shared Cognition/Mental Model Congruency Measure. See 3.3.1 Methodology Ft. Leavenworth.
- Created a goal directed cognitive task analysis of for the Joint Personnel Recovery Agency. See 3.3.3 Methodology Joint Personnel Recovery Agency
- Collected workload, social network and SAGAT data at Joint Personnel Recovery Agency. See 3.3.5 Methodology Joint Personnel Recovery Agency and Appendix A.
- Validated three separate SA models. See 5.0 Task 4 – Perform Social Network Analysis.
- Obtained remarkably high prediction scores for an initial model attempt. See 5.0 Task 4 – Perform Social Network Analysis
- Identified other variables to be included in future research to improve our model prediction capabilities. See 7.0 Phase II Tasks and Objectives**.**

Summary of Our SA Measure
Key Features:

- High predictive validity
- Includes multiple factors
- Ease of adding additional factors
- Has a predictive capability
- Real-time measure
- Works for distributed teams
- Works for dynamic teams
- Works in large or small C^3I organizations
- Unobtrusive
- Provides diagnostic information
- Provides graphical feedback
- Robust measure
- Can be extended to include team and individual SA.

Figure 21. Key Features of Our SA Measure

7.0 Phase II Tasks and Objectives

Even with our successful results from our phase I research, we still have several areas that need to be improved and addressed during phase II. The efforts we wish to pursue in phase II are in three different areas: modeling, creating an SA feedback system, and creating an SA measurement system. These specific tasks for these areas are listed below.

<u>Modeling Efforts</u>
In future modeling efforts we hope to validate our model more fully and create a model with a higher r^2 than .78. By including other factors related to team and shared SA we should be able to make this objective. Some of the tasks we will accomplish are:

- Validating our shared cognition models against full SAGAT data.
- Validate our model against the Leader Situation Model and the SSA level III data collected from Ft. Knox.
- Validate the prediction capabilities of our model. Can we predict future SA?
- Adding other factors into the model to see if we can improve our predictability. Such factors include:
    - Task participation
    - Collaboration tools used
    - Team environment
    - Acknowledgement
    - Previous interaction history
- Creating models based on team SA.
- Creating models based on levels of SA (Endsley, 1995a).
- Creating an automated network collection tool (Software translation).

<u>Feedback System</u>
Just providing a measure of SA is not enough. The measure needs to be easily interpretable and diagnostic to the commander or team lead. A score of .75 does not mean anything out of context. We propose creating a real-time visualization tool that shows how the team is performing. Along those lines we wish to perform three tasks:

- Creating a shared SA feedback system that provides real-time visualization of organizational shared situation awareness.
- Include a diagnostic capability that can make organization design recommendations based upon the results.
- Address the question "how should a shared SA estimate be displayed to the leader?"

<u>SA Measurement System</u>
Lastly, we want to take the measures used to build our models and create a real-time SA measurement tool that is unobtrusive, diagnostics and can work with both large and small teams of individuals. We propose two task for this effort.

- Build a real-time collection tool.
    - This could be a packaged chat system.
- Validate this tool as a means of measuring SA.

8.0 Research Personnel

Personnel & Assignments	
SA TECHNOLOGIES	
Cheryl Bolstad	Principal Investigator
CARNEGIE MELLON UNIVERSITY	
Cleotilde Gonzalez	Senior Researcher
John Graham	LTC, US Army
Mike Schneider	Computer Programmer

Our Key Team Members Bring Extensive, Relevant Capabilities

- 100% military system experience
- 100% systems requirements analysis
- 100% system development backgrounds
- 75% cognitive engineering experience
- 50% are Ph.D.s
- 50% have advanced degrees

Cheryl A. Bolted, PH.D
SENIOR RESEARCH ASSOCIATE – SA TECHNOLOGIES

EDUCATION

Ph.D.	2001	Psychology	North Carolina State University
M.A.	1988	Cognitive Psychology,	Florida Atlantic University
B.A.	1986	Computer Applications in Psychology	University of Colorado

EXPERIENCE (Summary)

Cheryl Bolstad has a bachelor's degree in Computer Applications in Psychology and a Master's degree in Cognitive Psychology. She completed her doctoral degree in Psychology at North Carolina State University specializing in cognition and aging. She has ten years of experience as a human factors engineer in a wide variety of projects including data collection and analysis, test design, report writing, computer programming, system administration, and data base development and analysis. She performed one of the first studies to determine sources of individual differences in SA. She has worked extensively in user interface design, situation awareness research, tests and measurements, and training. Dr. Bolstad's recent research has focused on developing methods for supporting team situation awareness in distributed systems and developing training systems for supporting situation awareness. She has created numerous computer tools for supporting researchers, including applications to test cognitive abilities, testing of SA in a team task and computer based training modules for training SA.

EXPERIENCE (Work History)

- Senior Research Associate, SA Technologies — 2003 – Present
- Research Associate, SA Technologies — 1998 - 2003
- Research Assistant, North Carolina State University — 1994 - 1998
- Human Factors Specialist, Monterey Technologies, Inc. — 1992 - 1994
- Research Data Analyst, Blue Cross Blue Shield of North Carolina — 1991 - 1992
- Research Analyst/Engineer II, Northrop Aircraft Division — 1988 - 1991

PROFESSIONAL ACTIVITIES & HONORS

- Human Factors and Ergonomics Society, Member since 1984
- Human Factors and Ergonomics Society Technical Program Committee
- American Psychological Association (Division 20), Member since 1995
- Awarded John Oliver Cook Graduate Fellowship (1998)

SELECTED PUBLICATIONS

Bolstad, C. A., & Endsley, M. R. (2000). *The effect of task load and shared displays on team situation awareness.* Paper presented at the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society, San Diego, CA.

Bolstad, C. A., & Endsley, M. R. (2003). *Measuring Shared and Team Situation Awareness in the Army's Future Objective Force.* Paper presented at the Human Factors and Ergonomics 47th Annual Meeting, Denver, Colorado.

Bolstad, C. A., Riley, J. M., Jones, D. G., & Endsley, M. R. (2002). *Using goal directed task analysis with army brigade officer teams.* Paper presented at the Human Factors and Ergonomics Society 46th Annual Meeting, Baltimore, MD.

Bolstad, C. A., Schneider, M., Graham, J., & Gonzelez, C. (2004). *Collaborative Tools and Shared Mental Models.* Paper presented at the Human Factors and Ergonomics Society 48th Annual Meeting, New Orleans, LA.

Endsley, M. R., & Bolstad, C. A. (1993). Human capabilities and limitations in situation awareness. In *Combat Automation for Airborne Weapon Systems: Man/Machine Interface Trends and Technologies (AGARD-CP-520)* (pp. 19/11-19/10). Neuilly Sur Seine, France: NATO - AGARD.

Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot situation awareness. *International Journal of Aviation Psychology, 4*(3), 241-264.

Cleotilde Gonzalez, Ph.D.
Director, Dynamic Decision Making Laboratory
Assistant Professor of Information Systems

EDUCATION

BSc	1986	Computer Science	Universidad de las Americas (Mexico)
MBA	1990	Business Administration	Universidad de las Americas (Mexico)
MSc	1992	Management Information Systems	Texas Tech University
PHD	1996	Management Information Systems	Texas Tech University

EXPERIENCE (SUMMARY)

Dr. Gonzalez's research has focused on dynamic decision making (DDM). In particular, she has focused in the psychology of decision making in complex, dynamic, situations. She created the Dynamic Decision Making Laboratory (DDMLab) at Carnegie Mellon University, where she is currently an Assistant Professor. Current research in the DDMLab includes projects aimed to investigate how individuals make decisions and improve their decision making processes. The DDMLab performs cognitive and computational modeling, and accomplishes laboratory studies to investigate the effect of several environmental and cognitive factors on the process of learning. Projects include the development of ACT-R cognitive models of situation awareness and learning in command and control during the execution of a battle. Other projects investigate the basic principles of automaticity in DDM.

EXPERIENCE (Work History)

- Systems Analyst, Universidad de las Americas, Mexico 1986
- Chief of department, Analysis and Programming 1987-1990

- Research Assistant, Texas Tech University 1991-1996
- Associate Professor, Computer Science, Universidad de las Americas 1996-1997
- Post-Doctoral Fellow, Graduate school or Industrial Administration, 1997-2000
- Carnegie Mellon University
- Assistant Professor of Information Systems, Social and Decision Sciences, Carnegie Mellon University 2000-current

PROFESSIONAL ACTIVITIES & HONORS

- Fellow – Human-Computer Interaction group (1993- present), Association of Computing Machinery
- Organizer of the Latin American Workshop at CHI-2001
- Chair, First Latin American Conference in Human-Computer Interaction (CLICH, 2003). Rio de Janeiro, Brazil.
- Fellow - Human Factors and Ergonomics Society. Cognitive Engineering and Decision Making Technical Group (1996 - present).

SELECTED PUBLICATIONS

Gonzalez, C. and Quesada, J. (accepted). Learning in Dynamic Decision Making: The Recognition Process. *Computational and Mathematical Organization Theory.*

Gonzalez, C. (accepted). Learning to Make Decisions in Dynamic Environments: Effects of Time Constraints and Cognitive Abilities. *Human Factors.*

Gonzalez, C., Lerch, F. J., & Lebiere, C. (2003). Instance-Based Learning in Dynamic Decision Making. *Cognitive Science, 27*, 591-635.

Gonzalez C. and Kasper G. Animation in User Interfaces Designed for Decision Support Systems: The Effects of Image Abstraction, Transition, and Interactivity on Decision Quality. *Decision Sciences Journal*. Volume 28. Number 4. Fall, 1997 (pp. 793-823).

John M. Graham
LT. COLONEL – US ARMY

EDUCATION

BS	1987	Military Science	United States Military Academy
MS	1997	Cognitive Systems Engineering	The Ohio State University

EXPERIENCE (Summary)

John Graham has been an officer in the U.S. Army for fifteen years. Since assignment, in 1993, as an Assistant Professor to the Army's Engineering Psychology Laboratory and Program, he has focused his career in the area of Decision-Making from a Human Factors and Human Computer Interaction approach. He has conducted and continues to conduct research on Army Tactical Operations Center (TOC) design and organization, Army Division-level Planning, and NORAD-Cheyenne Mountain Command Center Team Performance and Distributed Work. He has experience in Cognitive Task Analysis. John has working experience as a Scout Platoon Leader, an Airborne (parachute) Company Commander and Operations Officer, as well as a Space Operations Detachment Commander during a recent deployment in support of Operation Enduring Freedom. While at the United States Military Academy, he developed interactive class instruction in Human Error, Human-Computer Interaction, and Cognitive Psychology. He has served as the Director, of Space Operations Training; Instructor, Cheyenne Mountain Command Center; and as a certified Missile Officer in the Command Center. While in the Space Command, John lead the approval of the first study conducted in Cheyenne Mountain in fifteen years. Currently, John is assigned to Carnegie Mellon University to conduct research in Command and Control.

EXPERIENCE (Work History)

- PhD Student, Human-Computer Interaction Institute
  School of Computer Science — 2002 - Present
- US Army Officer (Space Operations) — 1987 - Present

SELECTED PUBLICATIONS

Graham, J. M., Gonzalez, C, Doyle, M. (2002) Using communication patterns in the design of an adaptive organizational structure for command and control. *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society*. HFES, California

Shattuck, L., Graham, J. M., Merlo, J., Hah, S. A. (2000) Cognitive Integration: A Study of How Decision-makers Construct Understanding in Evolving Contexts *Proceedings of the 44th Annual Meeting of the Human Factors and Ergonomics Society.* HFES, California.

Graham, J.M., Shattuck, L., Shevlin, J. S., Merlo, J., Hah, S. A. (1999) Analysis of Information Contained in Division-Level Operations Orders *STRICOM/ DARPA Tech Report* Orlando, Florida.

Woods, D.D., Watts, J.C., Graham, J.M., and Smith, P.J. (1996)Teaching Cognitive Systems Engineering. *Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society*, HFES, California.

Mike Schneider
SENIOR RESEARCH PROGRAMMER
HUMAN-COMPUTER INTERACTION INSTITUTE
CARNEGIE MELLON UNIVERSITY

EDUCATION

B.S.	1999	Computer Science	University of Oklahoma
M.S.	2001	Human-Computer Interaction	Carnegie Mellon University

EXPERIENCE (SUMMARY)

Mike Schneider has a bachelor's degree in computer science from the University of Oklahoma, and a Master's degree in Human-Computer Interaction from Carnegie Mellon University. Mike has a passion for creating software that is easy to use and gets the job done. His research interests currently focus on using social network analysis techniques to understand the organizational issues of military transformation. He is also interested in ways of visualizing large social network datasets, and in techniques for providing real-time displays of social and organizational data to increase situation awareness in military settings.

SELECTED PUBLICATIONS

Manalavan, P., Samar, A., Schneider, M., Kiesler, S., & Siewiorek, D. (2002). In-car cell phone use: Mitigating risk by signaling remote callers, CHI 2002 extended abstracts on Human factors in computing systems, Minneapolis, April 22-25.

Baker R.S., Corbett A.T., Koedinger K.R., Schneider, M.P. (2003). A Formative Evaluation of a Tutor for Scatterplot Generation: Evidence on Difficulty Factors. Proceedings of the Conference on Artificial Intelligence in Education, 107-115.

Schneider, M., Graham, J. Bauer, A. Bessiere, K. Gonzalez, C. "Shared Mental Models in Military Command and Control Organizations: Effect of Social Network Distance." Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, New Orleans, LA (9/1/2004)

Schneider, M, and Kiesler, S. "Calling While Driving: Effects of a Remote Traffic Display". CHI 2004 conference on Human factors in computing systems. (submitted)

Schneider, M. and Graham, J. "Effect of External Memory Tools on Task Performance in a Future Command and Control Environment: A Virtual Experiment using Construct". North American Association for Computational Social and Organizational Science (NAACSOS) 2004 Annual Proceedings.

9.0 References

Argote, L. (1999). *Organizational Learning: Creating, Retaining and Transferring Knowledge*: Kluwer Academic Publishers.

Bolstad, C. A., & Endsley, M. R. (1999). *Shared mental models and shared displays: An empirical evaluation of team performance.* Paper presented at the 43rd Annual Meeting of the Human Factors and Ergonomics Society, Houston, TX.

Bolstad, C. A., & Endsley, M. R. (2000). *The effect of task load and shared displays on team situation awareness.* Paper presented at the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society, San Diego, CA.

Borgatti, S. (2002). *Basic Social Network Concepts.* Paper presented at the AoMPDW, Denver.

Borgatti, S. P. (1994). A quorum of graph theoretic concepts. *Connections, 17*(1), 47-49.

Brannick, M. T., Roach, R. M., & Salas, E. (1993). Understanding team performance: A multimethod study. *Human Performance, 6*(4), 287-308.

Brass, D. J. (1995). Creativity: It's all in your social network. In C. M. Ford & D. A. Gioia (Eds.), *Creative action in organizations* (pp. 94-99). London: Sage.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in experts team decision making. In N. J. Castellan Jr. (Ed.), *Current Issues in Individual and Group Decision Making* (pp. 221-246). Hillsdale, NJ: Erlbaum.

Carley, K. M. (2003). *Dynamic Network Analysis.* Paper presented at the Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers.

Carley, K. M., Ren, Y., & Krackhardt, D. (2000). *Measuring and modeling changes in C^3I architectures.* Paper presented at the 2000 Command and Control Research and Technology Symposium, Naval Postgraduate School, Monterey, CA.

Dekker, A. (2002). Applying social network analysis concepts to military C4ISR architecturs. *Connections, 24*(3), 93-103.

Dominguez, C. (1994). Can SA be defined? In M. Vidulich, C. Dominguez, E. Vogel & G. McMillan (Eds.), *Situation awareness: Papers and annotated bibliography (AL/CF-TR-1994-0085)* (pp. 5-15). Wright-Patterson AFB, OH: Armstrong Laboratory.

Endsley, M. R. (1993). A Survey of situation awareness requirements in air-to-air combat fighters. *International Journal of Aviation Psychology, 3*(2), 157-168.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37*(1), 65-84.

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

Endsley, M. R., & Jones, W. M. (1997). *Situation awareness, information dominance, and information warfare* (No. AL/CF-TR-1997-0156). Wright-Patterson AFB, OH: United States Air Force Armstrong Laboratory.

Entin, E. E. (1999). *Optimized command and control architectures for improved process and performance.* Paper presented at the Command and Control Research and Technology Symposium, Newport, RI.

Entin, E. E., & Entin, E. B. (2001, June 19-21). *Measures for Evaluation of Team Processes and Performance in Experiments and Exercises.* Paper presented at the 6th International CCRTS, Annapolis, Maryland.

Entin, E. E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors, 41*(2), 312-325.

Espinosa, J. A., Kraut, R. E., Slaughter, S. A., Lerch, J. F., Herbsleb, J. D., & Mockus, A. (2001). Shared Mental Models and Coordination in Large-Scale, Distributed Software Development. *ICIS 2001*, 513-518.

Fjelde, K. M., & Switzer, F. S. (1994). Placement of crew's communication and coordination training in process control. *Perceptual and Motor Skills, 18*(527-551).

French, H., & Hutchinson, A. (2002). *Measurement of Situation Awareness in a C4ISR Environment.* Paper presented at the 7th International Command and Control Research and Technology Symposium.

Graham, J. M., Schneider, M., & Gonzalez, C. (2004). *Report Social Network Analysis of Unit of Action Battle Laboratory Simulations: Ft Leavenworth 2003 & Ft Lee 2004* (No. DDML Technical Report CMU-SDS-DDML-04-01): Carnegie Mellon University, Social & Decision Sciences, Dynamic Decision Making Laboratory.
Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: North-Holland.
Krackhardt, D. (1994). Graph Theoretical Dimensions of Informal Organizations. In K. M. Carley & M. J. Pretula (Eds.), *Computational Organization Theory* (pp. 89-111). Hillsdale, NJ: Lawrence Erlbaum.
Krackhardt, D., & Hanson, J. (1993). Informal Networks: The Company Behind the Chart. *Harvard Business Review, 71,* 104-111.
Monge, P. R., & Contractor, N. (1988). Communication networks: Measurement techniques. In C. H. Tardy (Ed.), *A handbook for the study of human communication* (pp. 107-138). Norwood, NJ: Ablex.
Mosier, K. L., & Chidester, T. R. (1991). Situation assessment and situation awareness in a team setting. In Y. Queinnec & F. Daniellou (Eds.), *Designing for Everyone* (pp. 798-800). London: Taylor and Francis.
Rouse, W. B., Cannon-Bowers, J. A., & Salas, E. (1992). The role of mental models in team performance in complex systems. *IEEE transactions on systems, man, and Cybernetics, 22*(6), 1296-1308.
Salas, E., Prince, C., Baker, D. P., & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human Factors, 37*(1), 123-136.
Salas, E., Rozell, D., Driskell, J. D., & Mullen, B. (1999). The effect of team building on performance: An integration. *Small Group Research, 30*, 309-329.
Scott, J. (1992). *Social Network Analysis*. Newbury Park, CA: Sage Press.
Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors, 41*(1), 61-71.
Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
Williges, R. S., Johnston, W. A., & Briggs, G. E. (1966). Role of verbal communication in teamwork. *Journal of Applied Psychology, 50*, 473-478.

Appendix A: Joint Personnel Recovery Center SAGAT Queries

1. How many isolated incidents are you aware of? # _________

2. How many of these isolated incidents have been verified and validated as actual incidents? #_________

3. Who is the SMC (SAR Mission Coordinator) for each incident?

(1)____________________	(5)____________________	(9)____________________
(2)____________________	(6)____________________	(10)___________________
(3)____________________	(7)____________________	(11)___________________
(4)____________________	(8)____________________	(12)___________________

4. Indicate the **number and status** of the isolated personnel (IP) for each incident

(1) #____OK	(5) #____OK	(9) #____OK
#____Slightly injured	#____Slightly injured	#____Slightly injured
#____Severely injured	#____Severely injured	#____Severely injured
(2) #____OK	(6) #____OK	(10) #____OK
#____Slightly injured	#____Slightly injured	#____Slightly injured
#____Severely injured	#____Severely injured	#____Severely injured
(3) #____OK	(7) #____OK	(11) #____OK
#____Slightly injured	#____Slightly injured	#____Slightly injured
#____Severely injured	#____Severely injured	#____Severely injured
(4) #____OK	(8) #____OK	(12) #____OK
#____Slightly injured	#____Slightly injured	#____Slightly injured
#____Severely injured	#____Severely injured	#____Severely injured

5. What is the current tactical situation around the IPs for each incident? (check one)

(1)	____ High threat ____ Medium threat ____ Low threat	(5)	____ High threat ____ Medium threat ____ Low threat	(9)	____ High threat ____ Medium threat ____ Low threat
(2)	____High threat ____ Medium threat ____ Low threat	(6)	____High threat ____ Medium threat ____ Low threat	(10)	____High threat ____ Medium threat ____ Low threat
(3)	____High threat ____ Medium threat ____ Low threat	(7)	____High threat ____ Medium threat ____ Low threat	(11)	____High threat ____ Medium threat ____ Low threat
(4)	____High threat ____ Medium threat ____ Low threat	(8)	____High threat ____ Medium threat ____ Low threat	(12)	____High threat ____ Medium threat ____ Low threat

6. What appropriate JTF and subordinate staff sections are aware of this incident?

(1) ____________________ ____________________	(5)____________________ ____________________	(9)____________________ ____________________
(2) ____________________ ____________________	(6)____________________ ____________________	(10)____________________ ____________________
(3) ____________________ ____________________	(7)____________________ ____________________	(11)____________________ ____________________
(4) ____________________ ____________________	(8)____________________ ____________________	(12)____________________ ____________________

7. What additional assets do you require to conduct a recovery?

(1) ____________________ ____________________	(5)____________________ ____________________	(9)____________________ ____________________
(2) ____________________ ____________________	(6)____________________ ____________________	(10)____________________ ____________________
(3) ____________________ ____________________	(7)____________________ ____________________	(11)____________________ ____________________
(4) ____________________ ____________________	(8)____________________ ____________________	(12)____________________ ____________________

Appendix B: Validation of Our Communication Measure

Introduction
The data collected here was used to validate our communication measure. We wanted to verify that self-reported communication measures were as effective at measuring communication between individuals as actual chat data collected during the exercise.

Methodology

In February 2004, the Combat Service Support Battle Laboratory (CSSBL) gathered twenty-eight experienced army officers to serve as participants in a prototype Unit of Action logistics command and control staff. Each participant was assigned to a cell with two to five other participants. The participants gathered information, coordinated with appropriate staff members, and entered battlefield actions into the simulation. Observations and data collection were conducted over two days immediately following a two-week training period. A plan-execute-plan-execute cycle was used.

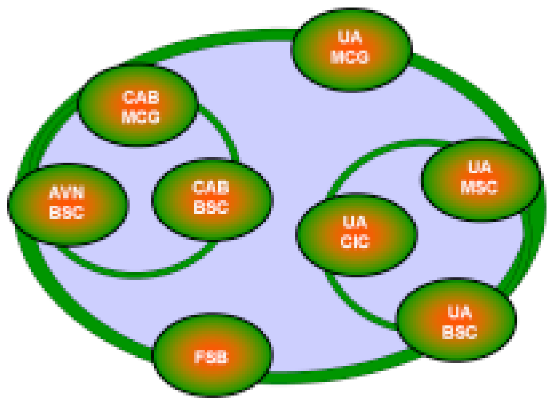
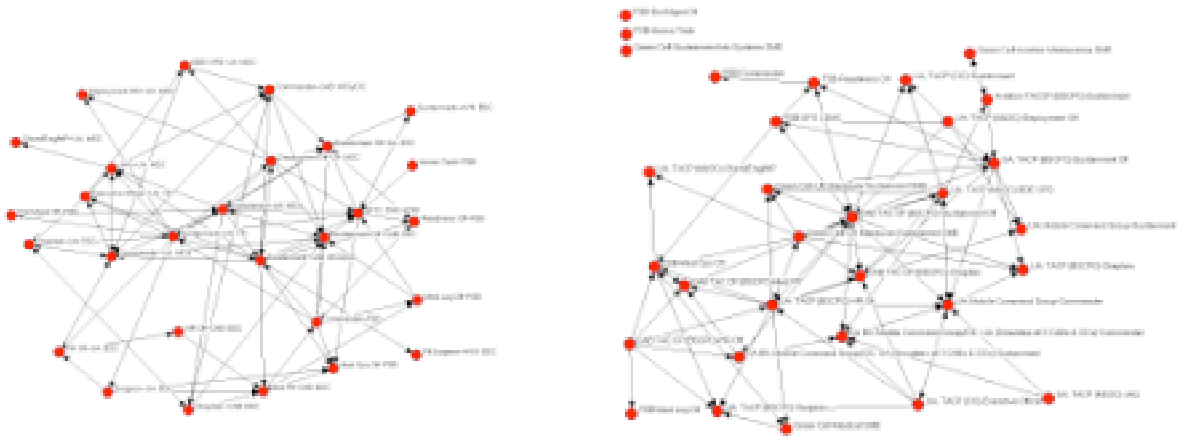


Figure 1B. Cellular Structure of CSSBL Experimental Network Organization.

Data

Self-report data was collected every 60-90 minutes using a networked questionnaire. The questionnaire asked participants to report people they had communicated with during the time since the previous questionnaire. They could give up to 10 responses by selecting participants from pull-down menus. The responses were ordered by the frequency of communication during the previous session. Questionnaire data was collected for a total of 6 sessions. Network graphs were constructed in ORA and analyzed. Figure 1A is a visualization for the self-report based network graph of the CSSBL participants for collection session 3.

Concurrently, chat network data was collected throughout the experiment. Each workstation was assigned to a specific user and their input to the chat room was continuously time-stamped and logged with list a of recipients on the individual workstations. A script on the server periodically, uploaded the chat log from each workstation, and created a time based chat log for the entire network. Figure 2A is a visualization derived from summing all chat conducted during the collection session 3.



A. B.

Figure 2B: Representational Network Graph

Figure 2B shows the representational network graphs of a self-report communications questionnaire versus a chat log derived network graph. a. self-report communications data for a ninety-minute collaboration session, b. chat data summed over the same ninety-minute collaboration session.

During this experiment there were multiple methods for participants to collaborate. Chat was only one of the available methods. Logging only one method of collaboration missed the other communications reported in the self-report questionnaire. Figure 2A demonstrates the missing data problem in that the self-reported graph (a) is fully connected while the chat data (b) has three nodes (upper left) that are disconnected.

To understand how much of the communications are missed by logging only one communications channel, we constructed a time series graph of network density for the chat data and the self-report data (Figure 3A). Network density served as a gross-level representation of the connectedness of a graph. Network density is the number of actual links observed between the members of an organization divided by the number of all possible links between the members of an organization (Freeman, 1979). A fully dense network/organization would have every person (node) linked to every other person.

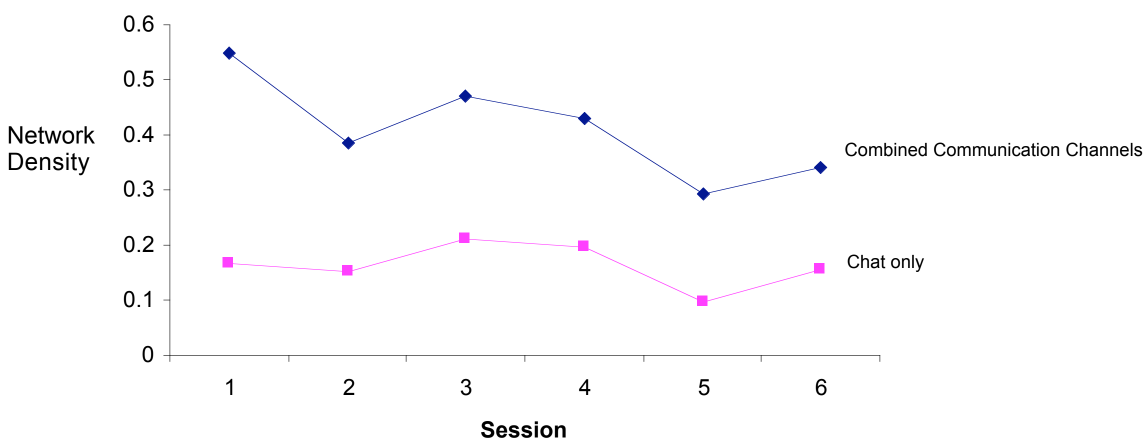


Figure 3B. Comparison Between Network Density

Figure 3B shows the comparison Between Network Density of Self-Report Communications and logged chat Communications. Density is the actual communication links divided by the possible communication links in a 28 member organization.

Results

Overall, average self-report network density is 41% while average chat network density is 16%. Assuming that the self-report network density represents communication over all potential collaboration channels, then using chat log only data would miss a significant portion of the communications network. However, the chat network graph followed the same network density pattern as the self report network graph but at a factor of ~1.56 less. Based on this data and given that the chat is only one of a subset of the communications mediums available (face-to-face, chat, voice-over-IP, instant messaging, or email), our use of the self-report is more representative of combined communication medium quantity. Further, any shared situation solutions must account for all communication mediums.

Appendix C: Modeling Process

Baseline Model (Model 1)

As a baseline for model efforts we used a simple model that states that SA is directly measured by communication. The baseline variable used in this model is direct communication. In essence this is self-reported measure of who reported talking to whom.

$$D_{ij} : DirectCommunication$$

During each iteration we tracked direct communications and used it as a performance baseline for our metric development. Direct communication between organizational members has been repeatedly demonstrated as a key variable in the development of shared mental models, shared situation awareness, and transactive memory. Many approaches to shared mental model and shared situation awareness rely solely upon direct communication to estimate the 'sharedness' between two organizational members.

$$SSA_{ij} = \square D_{ij}$$

As such a simple formula that considers direct communication in a linear relationship to shared situation awareness is an excellent performance baseline. For each metric iteration, we calculate direct communication based shared situation model estimation and compare against the performance of our metric. In terms of the algorithm, shared situation awareness (SSA) between two organizational members (i,j) is a function of whether or not they have directly communicated (Dij) during the time period of interest. Our goal in using this as a benchmark is to progressively beat the results achieved with a more informed Shared Situation Awareness metric.

Situation Model Metric

Our first effort went to estimating the situation model. Endsley (1995b) describes the situation model as "a schema depicting the current state of the mental model of the system" that is dynamic and constantly updating based on the evolving context. The model incorporates the value of different system parameters and includes and understanding of the dynamics of the system.

Since the ultimate goal of our research is to develop a model of shared situation awareness between members of an organization, the situation model we wanted to understand is that of the organization itself. The situation model 'system' for the purposes of this metric is the command and control organization providing planning and execution support to a military operation. The military command and control organization is a system of people and collaborative tools that exists in a dynamic environment. The organization can be subject to system-wide or specific component surges in activity. The organization members must adjust their behaviors by the correct perception and comprehension of the other components of the organization.

Model 2

We developed a hypothesized metric of situation model after an extensive review of the literature. Our original literature review and experience indicated that an individual's situation model of an organization is a function of physical proximity, network distance (nodes on the geodesic) &

organization communication status (authoritativeness). Following a description of the elements, they will be algorithmically integrated to produce a measure of Organization Situation Model.

One critical component of organization situation model is a mental model of roles and tasks (Entin & Serfaty, 1999). A organization mental model is an accurate understanding of who is responsible for what tasks and what the information requirements are for the tasks (Cannon-Bowers, Salas, & Converse, 1993; Rouse, Cannon-Bowers, & Salas, 1992). Organization situation models include transactive memory which is knowing, in a group or organization, who knows what information (Argote, 1999). Organization situation models can be trained/achieved via rehearsal (mental or actual) and cross-training. Measurement in the lab and in the field can be done through instruments such as the SAGAT (Endsley, 1995a) or workload estimation (Entin, 1999).

$$P_{ij} : Physical\,\mathrm{Pr}\,oximity$$

Physical proximity has been found to favor the development of models of others and improves performance. In two studies, Bolstad & Endsley (1999; 2000) found that collocation or proximity allows observation of another's activities. Through this observation, team members are able to more accurately obtain information about other's capabilities, tasks, and situation and are better able to establish and maintain a situation model of the people they interact with. In the case of Graham, Schneider, and Gonzalez (2004), we found that physical collocation was twice as likely to produce a shared mental model (figure 1C).

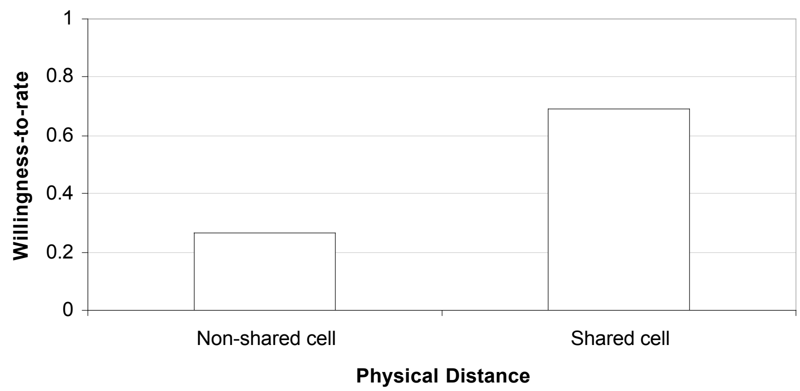


Figure 1C: Mean Shared Mental Model By Physical Distance
(shared cell vs non-shared cell) (n = 3028) (Graham et al., 2004)

We measure physical distance based on the metric distance between individuals (i,j) in the organization. If two members are physically collocated, we consider this a distance of zero. As they become more geographically dispersed, so does physical distance.

$$C_{ij} : Geodesic(SocialNetworkDist)$$

Multiple studies have also found that communication supports situation model development (Salas et al., 1999). Team members that communicate directly communicate tend to understand each others tasks and situation and are able to gather information about the other's capabilities (Graham et al., 2004).

We extend the definition of communications beyond direct communications to include the chain of communication in terms of the number of nodes on the geodesic between two agents. Krackhardt & Hansen (1993) found that, as communication network increased, knowledge about an organizational member decreased. This communication network distance is independent of physical distance in an organization.

We measure network distance based on the number of edges in the geodesic between two members of the organization. The geodesic is the shortest number of edges between two members (i, j). An edge is a communication link between two members of the organization. Even if two organizational members do not directly communicate, there is a likely set of communication links with other members that will connect them.

$$A_{ij} : NetworkLeadership$$

Members in close proximity to a leader are in the military C2 culture, more likely to have a good situation model of their leader than other organizational members at an equal distance. This phenomenon occurs because, in the military C2 culture, leaders are expected to have the most correct situation awareness (French & Hutchinson, 2002) and explicitly state their assessment of the situation and provide their intent for future activities (FM 100) to their immediate leaders and subordinates. The assumption is that the subordinates have the opportunity to develop better situation models of leaders because the leaders are explicitly stating and updating their situation models.

Leaders can be identified using two different constructs. The first is the formal leadership position. The formal leadership position is described by the published organizational chart. However, Krackhardt (1994) has found that this misses a large percentage of the people functioning as leaders without a formally assigned leadership role. A network-based description of informal leadership accounts for all organizational members, whether assigned a leadership role or not, that function as organizational leaders.

Network-based informal leadership is measured through the eigenvector centrality in the dynamic communications network (Scott, 1992). This descriptor of leadership assigns members with higher eigenvector centrality as leaders of leaders, and members lowest in eigenvector centrality hold strictly subordinate roles.

$$SituationModel_{ij} = \square A_{ij} + \square P_{ij} + \square C_{ij}$$

We took a very simplistic approach to the metric during this first iteration and created an additive algorithm. Leadership, physical proximity, and geodesic are each multiplied by a context determined constant and added to produce an estimate of an individual's situation model. Based on this algorithm, an individual will have their best situation model of leaders that are physically collocated and in direct communication. Organizational members will have the poorest situation model of non-leaders that are geographically distant and are high in geodesic edge count of the communications network.

<u>Model 2 Experimental Validation</u>

The situation model algorithm was validated on a data set collected at an organization experiment conducted at Ft Leavenworth, Kansas. The data set is from a trial fifty-six army officer organization. Each officer served as role-players for an experimental command and control staff that was put through a computer scenario for four days. The role-players were assigned to a functional cell with three to eight other role-players. The role-players gathered information, coordinated with appropriate staff members, and entered battlefield actions into the simulation. Partitions or walls separated the five cells, so that a participant could talk directly to members of his own cell, but could only communicate with members of other cells using the communication tools. A plan-execute-plan-execute cycle was used in the scenario.

While we would have preferred to log actual communications regardless of medium, the computer system for this particular experiment would not support an automated logger. As a result, a self-report questionnaire was employed as a simplified method for collecting all potential communications between members regardless of the communication method used. In a different two day experiment on twenty-eight role-players, we could log chat room activity and employed a self-report communications questionnaire (see previous task). While both the chat and the self-report data exhibited the same temporal change in communication quantity, the amount reported in chat is less than the self-reported communications. Based on this data and given that the chat is only one of a subset of the communications mediums available (face-to-face, chat, voice-over-IP, instant messaging, or email), the self-report is more representative of combined communication medium quantity.

During the experiment, data was collected every 60-90 minutes using networked questionnaire that asked the participants to self-report their communications during the prior session. They could give up to 10 responses by selecting participants from pull-down menus. A maximum of ten responses is appropriate as only one of the fifty-six participants reported communicating with the maximum number possible, with average response rate of four. Questionnaire data was collected for a total of 16 sessions. Three sessions were discarded due to collection software problems.

<u>Situation Model Metric</u>

Network distance and network leadership was calculated using the communications network graph developed from the communications self-report for each collection session. Agent-agent physical proximity matrix was constructed from the location of each agent. If an agent could effectively view the activities of another agent, they were considered to be collocated and a 1 was entered into the matrix. 0 is entered otherwise. Network distance, network leadership, and physical distance were entered into the algorithm to estimate situation model accuracy.

<u>Situation Model Field Measure</u>

The actual organization situation model accuracy for each individual was field measured using role-player estimates of other organizational member workload. Application of the measure depends upon two strong assumptions. The first is that to accurately estimate another's workload requires knowledge about the other's role, knowledge about the other person, knowledge about the current situation, and how the confluence of role, person, and situation interact to produce their workload rating. The second is that an accurate estimation of other organizational member's workload is indicative of an understanding of the organizational system.

Workload was measured using the NASA TLX (Task Load Index) (Hart & Staveland, 1988) assessment consisting of six workload parameters on a Likert scale. As in Entin (1999), participants were asked to rate themselves as well as five other people randomly selected from the other participants. This allowed us to sample the situation model accuracy at multiple time periods throughout the scenario with short questionnaires.

When rating other people, the role-players had the option of selecting "Don't Know" for each of the seven questions. In a typical laboratory experiment consisting of college sophomores, "don't know" would not have been an option and the participants would be expected to make their best guess. However, working with experts requires different methods. Experts tend to know when they do not have the knowledge to accurately answer a question. Adhering to a traditional experimental design and forcing the experts to blindly guess would have created frustration and decreased response validity. The ‘don’t know’ option reduced frustration and increased instrument validity.

The situation model accuracy was determined by comparing workload estimation of a particular role-player against the self-reported workload. This measure was computed by summing the absolute differences between the ratee’s self-reported ratings and the rater’s estimations. Congruence scores could range from 0 (indicating perfect congruence) to 36 (indicating perfect incongruence). If a role-player chose ‘don’t know’ the situation model accuracy was assigned a score of zero.

Figure 2C is a graph of actual situation model accuracy over all experimental sessions. Note that the overall situation model accuracy mean initially decreases and then levels out. During the experiment, we observed that participants migrated to different collaborative tools and tool use may have implications to situation model accuracy. This phenomenon will be considered in development of the iteration #2 algorithm. Overall, it appears the mean situation model accuracy suffered a significant decrease over the life of the organization ($p<.01$, $F(34, 2128)=24.94$).

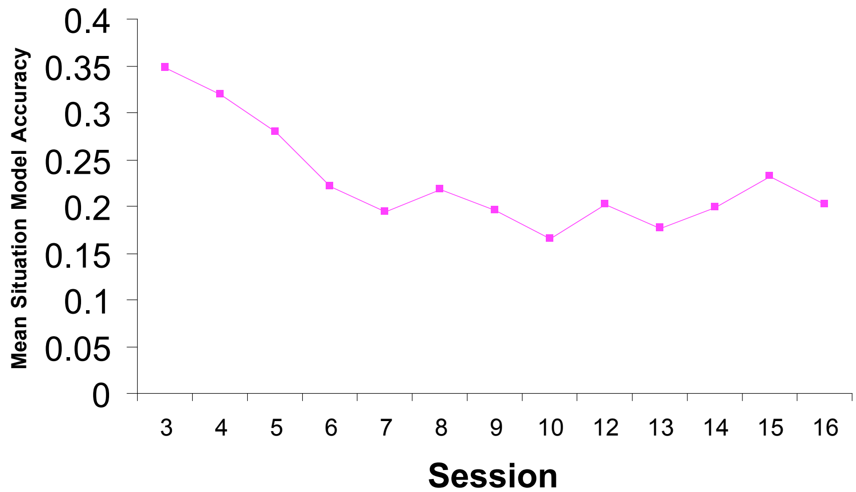


Figure 2C: Mean Situation Model Accuracy By Session

<u>Metric Validation</u>
To validate the algorithm situation model estimate against the field measured situation model we ran a multiple least squares regression on the variables of interest. For a point of comparison, we also calculated social network distance, physical distance (proximity), and authoritativeness as alternative surrogate measures of situation model accuracy. In the regression model, session was used to control for the effect of learning in this new organization. Regression results indicate that the situation model metric was the best predictor of field measured situation model accuracy.

Figure 3C is a graph of the situation model metric accuracy over the course of the experiment as compared to the mean performance of the simple direct communication measure (Dij) as a baseline for performance. The mean baseline performance (Dij) only accounted for 15% of the variance in field measured situation model accuracy during the experiment. The mean r-square for the situation model accuracy metric was .24 ($p<.001$; $F(3, 2298) = 564$). The metric performance steadily improved as applied to organizational data collected later in the experiment with its best performance accounting for 41% of the variance. The situation model metric clearly outperformed the baseline for metric performance

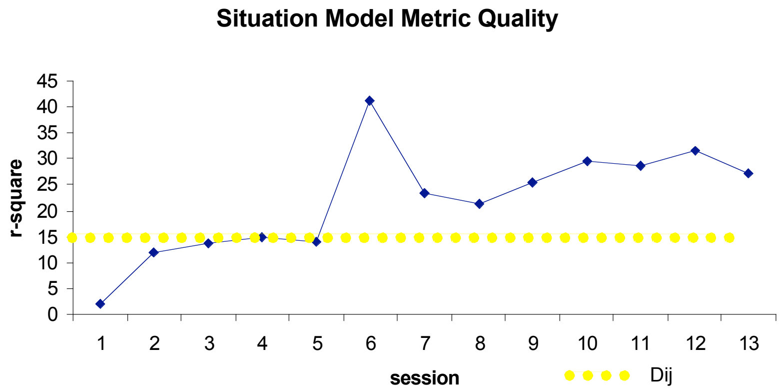


Figure 3C: Situation Model Metric Quality

Figure 3C is a graph of the quality of the shared situation model metric over the course of the experiment. Dij is the mean performance of the direct communication baseline metric for comparison.

Our situation model metric performed well. Any time a researcher finds a metric that accounts for 30-40% of the variance of any variable in a large organization is considered a publishable result. Further, we nearly doubled the performance over the baseline metric of direct communication. However, for a military real-time application, we need performance to be in the 70-80% accuracy range. The second iteration of our metric adjusted for the lessons learned and observations from the first metric iteration.

Model 3

Model 3 took advantage of observations from iteration #1 & designed a metric for shared situation awareness level III.

The third iteration of our metric took advantage of an observation made during the first iteration experimentation phase. Specifically we found that the organizational member made significantly more accurate workload estimates of organizational members with similar backgrounds as themselves ($p<.01$; $F(29, 1539)= 22.96$). Background similarity, in this case, considers years of service, branch of training, and types of staff experience/assignments.

$$H_{ij} : Homophilly$$

In the social network literature, background similarity has strong connections with the concept of homophilly. Homophilly theory states that members are more likely to create communication ties with other group members who they deem to be similar. In colloquial terms, "birds of a feather flock together." Brass (1995) observes that "similarity is thought to ease communication, increase predictability of behavior, and foster trust and reciprocity" (Monge & Contractor, 1988). Work by Espinosa, Slaughter, Herbsleb, Kraut, Lerch, and Mockus (2001) demonstrated that background familiarity improves the shared mental model between members of a team. In this case, we are not using homophilly to estimate the likelihood that two people will communicate, but instead we are seeking to estimate the shared situation awareness between two people in an organization. Hij represents a background similarity score between any two organizational members (i,j).

$$SharedSA_{ij} = \square A_{ij} + \square P_{ij} + \left(\frac{\square H_{ij}}{\square C_{ij}} \right)$$

Our the metric for shared situation awareness incorporated all of the variables that were used in the situation model metric. However we extended the model to account for the effects of homophilly. Our first iteration observations indicated that the effect of homophilly was greatest at shorter network distances than longer network distances. Therefore, homophilly was integrated into the shared situation awareness metric as a mediator of the effect of network distance. Based on this algorithm, two organizational members (i,j) are higher in shared situation awareness (Shared SA) if at least one of the members is high in network leadership, the two are physically collocated, they are in direct communication and they have similar backgrounds. Organizational members will have very low shared situation awareness if neither is a network leader, they are geographically dispersed, they are high in geodesic edge count of the communications network and they are low in background similarity.

Model 3 Experimental Validation

The shared situation awareness algorithm was validated against a data set collected at an organizational experiment conducted at Ft Knox, Kentucky. The data set is from a trial 256 member command and control organization. Similar to experiment #1, each participant served as role-players for an experimental command and control staff that was put through a computer scenario for eighteen days. The role-players were assigned to a functional cell with three to eight other role-players. The role-players gathered information, coordinated with appropriate staff members, and entered battlefield recommendations/decisions. The participants could communicate with their remote colleagues via email or radio network. During the exercise, participants completed an on-line survey. All answers were based on the time period since the last survey was collected. The survey was implemented as a web form, which the participants completed in an ordinary web browser. To reduce interruptions during the scenarios, all answers were multiple choice.

Social Network Data

During the exercise, participants were asked to list the top 7 people they communicated with (in descending order). This is the same method used in the Ft. Leavenworth exercise except only 7 people are selected as opposed to 10. The communication survey was filled out by all participants 2-4 times per day, depending upon the pace of the operation.

Shared SA Metric

Network distance and network leadership was calculated using the communications network graph developed from the communications self-report for each collection session. Agent-agent physical proximity matrix was constructed from the location of each agent. If an agent could effectively view the activities of another agent, they were considered to be collocated and a 0 was entered into the matrix. If the participants were located in adjacent cells, a one was entered into the matrix. If the participants were located in different cells and separate parts of the large experimental site warehouse, a two was entered into the physical distance matrix. Lastly, if the participants were located on different installations (Ft Knox, Ft Lee, Ft Leavenworth, Ft Sill, or Ft Rucker) a four was entered into the physical distance matrix. Network distance, network leadership, and physical distance were entered into the algorithm to estimate situation model accuracy. Homophilly was calculated based on a similarity score from background information the participants provided in their user profile.

Shared SA (level III) Field Measure

The best validation of our metric would be against a congruency in SAGAT scores between each participant. A SAGAT would require the participants t provide extensive information about their perceptions, comprehension, and projections relative to the current environment and situation. However, due to the size and pace of the exercise, we were unable to administer a full SAGAT at every collection period. We were, however, able to employ a SA congruency measure to account for Endsley's (1995b) level 3 SA or the projection of what will happen in the environment.

To find congruence in level 3 SA, participants were asked "What are the two top risks to this operation in the immediate future" at each stop. They could choose from a total of twenty-two choices that were divided into categories of Friendly, Enemy, and Environment. Using this data, we constructed a congruence score for all pairs of organizational members.

Figure 4C is the mean situation awareness (level 3) congruence for the organization. In the early stages of the experiment, there were tremendous fluctuations in congruence as the organization trained and the individuals learned their roles. Further, we noted that congruence shifted based on whether the organization was conducting split planning and execution operations or focusing only on planning or execution. This contextual factor observation will contribute to future metric development.

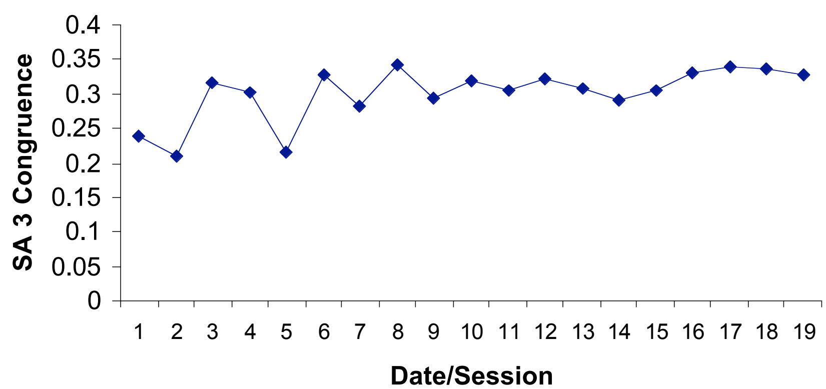


Figure 4C: Level 3 SA Congruency Scores

Figure 4C shows the mean SA level 3 congruence over the duration of the experiment (June, 04)

Metric Validation

To validate the algorithm situation model estimate against the field measured situation model we ran a multiple least squares regression on the variables of interest. For a point of comparison, we also calculated social network distance, physical distance (proximity), and authoritativeness, and homophilly as alternative surrogate measures of situation model accuracy. In the regression model, session was used to control for the effect of learning in this new organization. Regression results indicate that the shared situation awareness metric was the best predictor of field measured situation awareness (level 3) congruence.

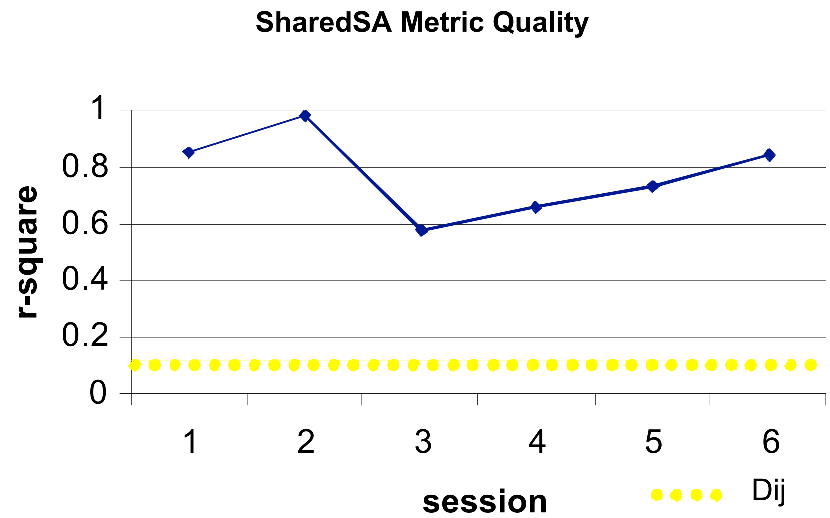


Figure 5C. Model 3 Validation Data

Figure 5C is a graph of the situation model metric accuracy over the course of the experiment as compared to the mean performance of the simple direct communication measure (Dij) as a baseline for performance. The mean baseline performance (Dij) only accounted for 15% of the variance in field measured situation model accuracy during the experiment. The mean r-square for the situation model accuracy metric was .24 ($p<.001$; $F(3, 2298) = 564$). The metric performance steadily improved as applied to organizational data collected later in the experiment with its best performance accounting for 41% of the variance. The situation model metric clearly outperformed the baseline for metric performance.

Overall, the shared situation awareness metric performed extremely well. Further, the metric is sufficiently valid for application to real-world shared SA tracking in military command and control organizations.