Dynamic Network Analysis of the Network-Centric Organization: Towards an Understanding of Cognition & Performance

Table of Contents

Appendices

Hourly Questionnaire

Dynamic Network Analysis of the Network-Centric Organization: Towards an Understanding of Cognition & Performance

Chapter 1

INTRODUCTION

As the network organizational form has proliferated, the requirement to describe, estimate and predict the dynamic structure as well as the need to measure performance in network-organizations has become critical (Carley, 2002b; 2003). Traditional, static organizational line charts no longer produce viable explanations of organizational behavior. As a network, the members are constantly reorganizing to the environment and therefore are rapidly changing the organizational structure (Graham, Schneider, Gonzalez, 2004).

The network-organization is becoming prevalent in command and control (C2) organizational design (Alberts et al, 1999; Evan, 1972). This research highlights two C2 network organizational contexts: terrorist organizations (Al Qaeda, Jemaah Islamiya, MILF, etc) and military command and control structures (Unit of Action, Unit of Employment, etc). Terrorists are often described as operating as a network-organization (Krebs, 2002; Lesser et al, 1999) and our own military is in the beginning steps to shed the old hierarchies and functional stovepipes for network-centric structure (Cebrowski & Garstka, 1998; Arquilla & Ronfeldt, 2000). Both examples share identities as organizations that are distributed over large distances, use multiple tools to support collaboration, are linked through a common identity and mission, and are required to flexibly reorganize to engage in problem-solving, decision-making and action (Ronfeldt & Arquilla, 2001).

Without some ability to understand the fluid organizational structure, the management of network-organizations is difficult at best. Management, leaders, and researchers need a method to understand the mediators of network change and evolution. Understanding and estimating behavior in a network organization requires an approach that departs from the traditional formal organizational chart and hierarchical reporting structures to account for the unique complexity of the organizational form. Social network theory is one such approach that provides a unique and powerful method understanding of the network organization. Further with the advent of today’s computing power many social network-based quantitative and qualitative analysis methods are now entering the mainstream management systems.

However, traditional social network analysis does not adequately address the dynamic nature of some organizational forms such as those operating in command and control environments. As a result, Dr Kathleen Carley (1998), Dr Jeffery Johnson (1998) and others have developed the concept of dynamic network analysis. Dynamic network analysis considers the factors that lead to temporal change in an organization’s network structure and performance.

While Dynamic Network Analysis is the right answer at the right time, Dynamic Network Theory is relatively young and unexplored. Relatively few studies and even fewer applications are available on Dynamic Network Analysis. This research proposes to extend the application of dynamic network analysis into the command and control operational environment and develop new Dynamic Network-based measures for the command and control context.

Dissertation Goal

I am interested in understanding how and why network organizations change from the perspective of Dynamic Network Analysis. Specifically, I intend to identify the factors that predict communication network change in a network organization and, given a change in the network, how the performance of the organization is altered. For instance military command and control organizations can self-organize into multi-functional teams/cells based on problem-solving requirements. This distributed sub-organizational team formation radically alters the communication network and overall organization structure. While it is understood that the current task load is one factor that explains team/cell membership and, therefore, communication network change, it is not clear what other factors contribute to the organizational shift.

The network changes I am specifically interested in happen in short periods of time. The rapidity of network change as indicated by the communications network is an organizational hallmark in terrorist and military command & control (C2) contexts (Ronfeldt & Arquilla, 2001). In both cases, wide ranging network change can occur in a matter of hours or days. In a terrorist organization, the launching of an operation or the capture of a member will cause a furious adjustment to the structure and performance of an organization (PACOM interviews, 2004). In a military command and control structure, an enemy engagement or resource criticality will cause the rapid shifting of problem-solving priorities and network structure (Unit of Action Operations & Organization Manual, 2002).

These rapid shifts in organizational structure have implications for organizational performance and measurement. Measures of organizational performance include situation model (Entin & Entin, 1999), mental model congruence (Entin & Serfaty, 2000; Graham, Schneider, Bauer, Bessiere, & Gonzalez, 2004), transactive memory (Carley & Ren, 2001), shared mental model (Salas et al, 1995) and shared situation awareness (Endsley &

Jones, 2001; 2002). Each of these measures have, as a common characteristic, the ability to describes how 'in-sync' the organizational membership is at a given point of time. Organizations operating in a fast-paced environment which are 'in-sync', tend to have higher levels of performance than those that are not (Salas et al, 1999; Arquilla & Ronfeldt, 2002)

To understand and describe rapid network change, I will use a combined approach consisting of experimentation and simulation. I will conduct multiple, large-scale experiments on military command and control network-organization structure in a pseudo-laboratory environment (Chapter 3). Using dynamic network analysis techniques, I will identify and characterize the factors that contribute to the change in structure of the organization(Chapter 4). Network change factors under consideration are homophily, physical distance, communication network distance, formal organizational structure, collaborative tool use, task & context.

These empirically based factors will be instantiated in new implementation of the proximity matrix in the ORA & DyNet models (Chapter 5)(Carley et al, 2003). DyNet currently relies upon the proximity matrix as a parameter to account for demographic and background similarity of network members. With an expanded implementation of the proximity matrix I will conduct multiple simulation experiments on terrorist network organizations for demonstration and validation purposes.

Lastly, the development of this research topic has generated strong interest in both the domain of military C2 and counter-terrorism analysts. As a result, use cases that support both domains have been constructed. Originally, I expected to purely focus my efforts towards applications oriented on understanding and supporting US Army command and control network organizations. However, the same research has implications for understanding and supporting intelligence analysts seeking to destabilize international terrorist network organizations. I have developed two use cases that would utilize network change modeling as well as the shared situation awareness metric (Chapters 4 & 5).

Motivation

From 2000-2002, I served as the Missile Officer in the Command Center of the Cheyenne Mountain Operations Center/ NORAD. While there I became increasingly aware of oversimplified explanations of organizational behavior. While there was a clear hierarchy and organizational line chart, I found myself completing tasks and missions using an informal organizational chart known only to myself. I realized then that even in strict hierarchies there are informal networks that do a lot of the real work. But this was hardly a laboratory setting and it was possible that my behaviors were not part of an organizational pattern, but was instead an individual anomaly.

However, in 2003, I was invited to observe and report on an experiment conducted at the Fort Leavenworth Battle Command Battle Laboratory. The US Army was in the opening phase of a ten-year organizational design process for a knowledge-centric command and control element. In support of this initial effort, the Fort Leavenworth Battle Command Battle Laboratory (BCBL) was conducting the first high fidelity experiment to determine organizational constructs that would support command and control in the Transformation Force. The experiment assumed a network-centric staff cell structure supported by a higher level of automation.

The following section is an explanation of this preliminary study and the findings that motivated my dissertation. I collated my observations into an ethnographic study that compared the envisioned communications performance against the actual communications performance of experienced military members role-playing command and control staff members. I developed the envisioned communications performance metric through a complex knowledge object-grouping task by an expert panel of military officers and Army scientists. These experts had extensive experience in the legacy force and tended impose a hierarchical organizational form.

The experimental measures of actual communication performance were derived from the data log of three offensive battle simulations conducted by the role-players over three days (8+ hours of coding). From this log, I could actually derive which role-player communicated with which other role-player during the course of the event. Given that all members had radios and computers that supported chat, there were no tangible limiting factors on who could speak with whom.

My goal in comparing the expected against the actual was intended to be a methods paper. I wanted to examine an alternative initial high-level quantitative approach to qualitative and low-level quantitative approaches. Qualitative approaches such as future incident forecasting (Smith et al, 1998) or process tracing methods (Woods, 1993) may be inappropriate because of a lack of specification of the organization concept. For the same reason, low-level quantitative measures such as those provided by organizational design systems (Entin & Entin, 2001) are also unsuitable.

Envisioned Communication Performance

The envisioned communication performance was developed long before the actual experiment. A representative set of Army intellectuals was assembled at Ft Leavenworth six months before the command and control experiment. Twenty-six Army officers ranging in rank from Brigadier General to Captain[1] served as participants in knowledge object (KO) development and grouping. All participants had worked with the military for a minimum of 7 years and the average time working with the Army was approximately 16 years.

Participants were given one day of military decision-making training (review) to provide common ground for discussion in future sessions. On day two, in four ninety minutes sessions, the group specified a separate set of KOs for each of the typical operational missions of an Army command and control element: Offense, Support, Defense, Stability. They then individually submitted ratings of each of the knowledge objects to specific operational missions as in Figure 1.1. On day three all participants were assembled and required to create “Natural Knowledge Clusters” based on mission-type for each of the knowledge objects.

---

[1] representing 13 military posts and 16 command organizations and nine Army and civilian scientists representing an equal number of theoretical approaches

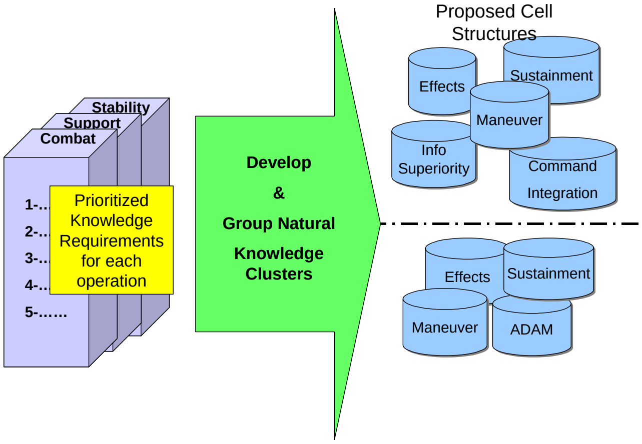


Figure 1.1. Knowledge Object development, prioritization and clustering resulted in cell structures designed for Combat (offensive, defensive), Support, and Stability operations.

Knowledge object ratings were distributed using a majority wins process. The "Natural Knowledge Clusters" resulted in a proposed cell structure for the experimental staff organization. The knowledge objects were then clustered a second time to maintain a consistent command and control organization across operational missions. The resulting experimental staff organization and the knowledge object distribution is represented in Figure 1.2.

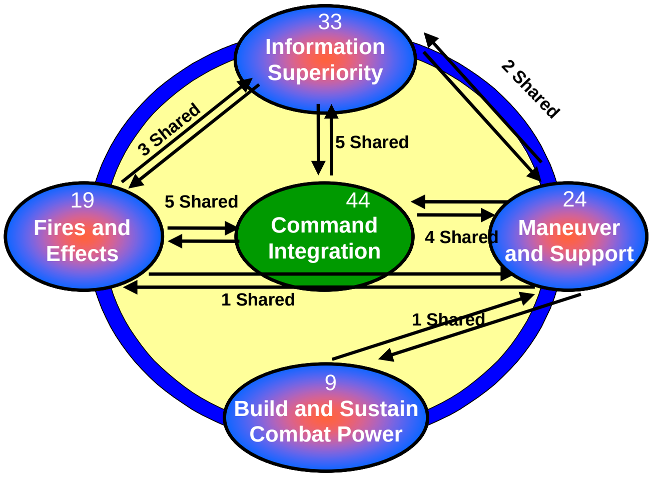


Figure 1.2. Cell structure resulting from initial Knowledge Object distribution workshop conducted at the Ft Leavenworth Battle Command Battle Laboratory. 111 of the knowledge objects are represented in this depiction.

Actual Communication Performance

In February 2003, the actual experiment using the knowledge object cell structure was conducted at the Fort Leavenworth BCBL. 26 Army officers served as role-players for the experimental command and control staff. They spent one week in team training. During this time, they were required to learn a) the concepts behind the experimental organization, b) a new method to make decisions in the experimental organization, c) their role in the structure of the experimental organization, and d) how to use the simulation software during the experiment.

The experiment itself was conducted as a command post exercise using OneSAF simulation software. The role-players gathered information and input actions on the battlefield via the simulation. Throughout the experiment, an average of fourteen data collectors entered observations, real-time, into a data log on Group Systems (Nunamaker et al, 1991). As a result every observable information transaction was captured. Each entered transaction was time-stamped with the participants, the discussion, and the outcome.

Findings

There are two major quantitative deltas in the actual vs. envisioned interactions. First, the amount of leader communication external to the staff organization was about 60% of his total communication. The distribution of communications to and from the commander position from the data log coding yielded an unexpected result. Specifically, about 60% of the commander's interactions were outside his command and control organization. Almost all of those interactions were to his subordinate leaders in the operation (Figure 1.3).

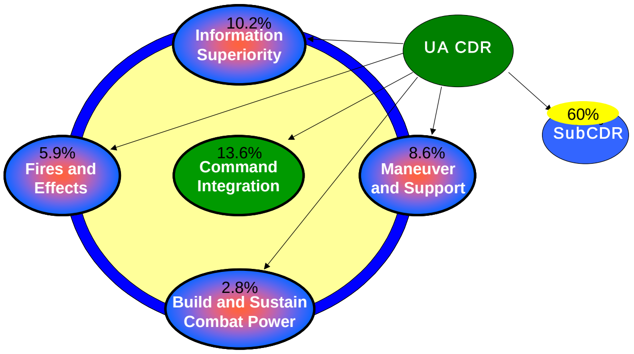


Figure 1.3. Envisioned (Desired) knowledge object distribution moderated communications compared to the actual leader (commander) communications during the first experiment

Second, the actual distribution of interaction within the staff did not match the envisioned distribution of interaction based on the knowledge object assignments (table 1.1). The envisioned or expected values had little relationship to the actual communication performance of the organization. The role-players were operating outside the expected organizational line chart.

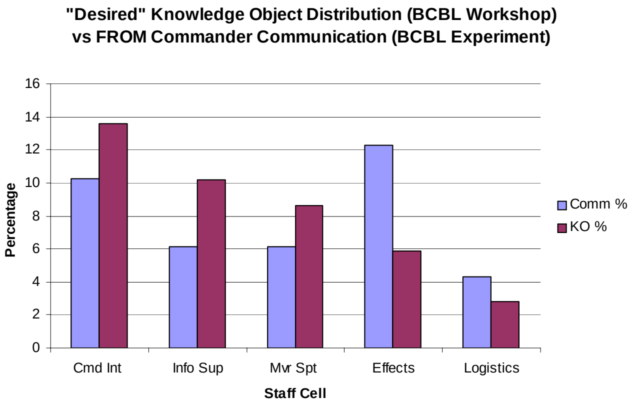


Table 1.1. Envisioned prioritization (KO) vs Actual (comm) communications per cell.

More important than the findings themselves was the reaction to the findings. The command and control practitioners and researchers were surprised! They had come to expect members of the military to use the hierarchical system so prevalent in the organizational design. They did not expect organizational members to form their own communication paths to complete work outside of doctrine. The role-player's behavior was considered unique.

However, stepping back from the command and control context, the findings should not have produced surprise. Both of these findings boil down to a simple concept: real-world behavior does not match the organizational chart. This same finding has been replicated by Krackhardt (1993) in his study of the informal organization. However, it has not been demonstrated in the context of a command and control environment.

Some of the practitioners stated that the organizational member's departure from the prescribed hierarchy might be more productive in the short run, but the chaotic process would lead to lower shared situation awareness in the organization overall. When I asked how they made this estimation, the replies referenced gut instinct and experience. None of the practitioners had a tangible measurement for shared situation awareness, their primary performance measure!

Conclusion

My observations and experience lead to a set of defined and tractable objectives for a dissertation. I wanted to find a usable network-based metric of organizational Shared Situation Awareness. However, the results have been much more far reaching. In the process of discovery, my work (along with the colleagues that participated) has changed one field of practice, enlightened a developing organizational design, informed decisions that are being implemented around the world, and instigated new research projects and lines of funding. At the core, the research process remained the same, but the findings & outcomes were much more diverse than expected. As a result, this dissertation includes topics that became relevant during the discovery process. While not core to SSA, they are core to the discovery process engaged in during this research.

REFERENCES

1.Bolstad, C. A., & Endsley, M. R. (2000) The effect of task load and shared displays on team situation awareness. In Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica, CA: HFES.

2.Bolstad, C. A., & Endsley, M. R. (1999**)** Shared mental models and shared displays: An empirical evaluation of team performance. In Proceedings of the 43rd Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica, CA: Human Factors and Ergonomics Society.

3.Endsley, M. R., and Robertson, M. M. (1996) Team situation awareness in aviation maintenance. Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society (pp. 1077-1081). Santa Monica, CA: Human Factors and Ergonomics Society.

4.Entin, E.E., Entin, E.B. (2001). Measures for Evaluation of Team Processes and Performance in Experiments and Exercises. In Proceedings of the 6th International Command and Control Research and Technology Symposium. June 19-21 2001, Annapolis, Maryland.

5.Gonzalez, C., Lerch, F. J., & Lebiere, C. (In press) Instance-Based Learning in Real-Time Dynamic Decision Making. Submitted to Cognitive Science.

6.Nunamaker, J., Dennis, A., Valacich, J., Vogel, D. & George, J. (1991) Electronic Meeting Systems to Support Group Work. Communications of the ACM, 7, July 1991, 40-61

7 . Smith, P.J., Woods, D.D., McCoy, E., Billings, C.E., Sarter, N.B., Denning, R., and Decker, S. (1998). Using Forecasts of Future Incidents to Evaluate Future ATM System Designs. ATC Quarterly, 6(1), 71-85

8.Shattuck L G, Graham J M, Merlo J L, Hah S (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. pp 1-116.

9.Vogel, Douglas R. & Nunamaker, Jay F. Jr. (1990, January). Group Decision Support system Impact: Multi-Methodological Exploration. Information & Management, 18, 15-28

10.Woods D.D. (1993) Process Tracing Methods for the Study of Cognition Outside of the Experimental Psychology Laboratory. In G. A. Klein, J. Orasanu and R. Calderwood, editors, Decision Making in Action: Models and Methods.

11.Woods, D.D. , Johannesen, L., Cook, R.I., and Sarter, N. (1994) Behind Human Error: Computers, Cognitive Systems, and Hindsight. Crew Systems Ergonomics Information and Analysis Center, Wright-A less tenuous organizational design will warrant qualitative and more specific quantitative approaches

12.Krackhardt, D., Hanson, J.R. (1993) Informal networks: The company behind the chart. Harvard Business Review, 71(4), 104-111.

Dynamic Network Analysis of the Network-Centric Organization: Towards an Understanding of Cognition & Performance

Chapter 6

Social Network Distance Matters: Social Awareness and Performance in the Command & Control Organization

This paper presents an investigation on the relationship between social network distance and social awareness in three large-scale experiments in military command and control. In the distributed environment, social awareness information is considered a prerequisite for effective collaboration, by communicating, planning and coordinating people's work activities (Dourish and Bellotti, 1992). Previous research has shown that physical distance is the gold standard for calibrating social awareness and obtaining high performance (Olson & Olson, 2000). However, social network distance may be equally or more important, as social network graphs inherently take into account the group's context and environment (Krackhart, 1994). We conducted this research on a series of large scale military experiments on command and control organization using computer-based collaborative tools as they engaged in a five day simulation exercise. As military command and control organizations are difficult to evaluate based on outcome and performance, we chose social awareness as a proxy. We hypothesized that in a distributed command and control organization, social network distance is a predictor of social awareness. We found that there is a very weak positive correlation between social network distance and physical distance, indicating that they are not the same construct. Further, we found that, controlling for physical distance; social network distance is a predictor of social awareness. This research provides a new construct for understanding and modeling social awareness and performance in large distributed command and control organizations.

INTRODUCTION

Olson & Olson (2000) found that physical distance matters to performance in group work. They stated that despite the gains achieved in distance communication during the last ten years by the internet and collaboration tools, performance in groups is still strongly predicted by physical distance. In this paper we argue that the Olson's focus on physical distance miss the effect of social network distance afforded by distance communication tools. This paper is based on our belief that social network distance is not the same as physical distance but is equally or more important to predicting performance. Social network distance is potentially more important because context and task requirements of a group are implicitly embedded within the social network graph (Krackhart, 1994).

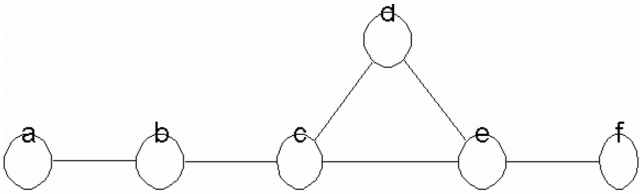


Figure 3.1. A simple network graph. (Borgatti, 1994)

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 (Figure 3.1) (Borgatti, 1994; Scott 1992). Communication data can be gathered by looking at various measures such as shared email headers or phone calls, or by surveying the individuals (Wasserman & Faust, 1994). Social network distance (often referred to as a geodesic) (Borgatti, 1994) is the number of links or actors between two members of a social network graph.

In a military command and control organization, measures of performance are not directly extractable from the end-state of decision-making. However, behavioral measures of constructs with proven correlation to good performance are possible. One such measure is Social Awareness. Social Awareness information is considered a prerequisite for people to obtain effective collaboration, by communicating, planning and co-ordinate their work activities (Dourish and Bellotti, 1992).

Dourish & Bellotti (1992) state the definition of Social Awareness as, as *“the understanding of the activity of the others, which provides a context of your own activity*". Alternatively, Gutwin and Greenberg (1996) use a tool-based definition, *the collection of up-to-the minute knowledge a person uses to capture another’s interaction with the workspace, e.g. where other participants are working, what they are doing, and what they have already done in the workspace.* Moran and Anderson (1990) drew a parallel to *peripheral awareness*, based on the human capacity to process peripherally non-attended aspects of others in the work environment. Lastly, Tollmar et al (1996), used *social awareness* to describe awareness about the social situation of the members.

Much of the previous research in this area uses questions about the team, task, and situation to generate a congruence measure for a group’s mental model (Espinosa et al, 2001; Cannon-Bowers & Salas, 1993; Entin & Serfaty, 1999). For instance, Entin (1999) found that a measurement of workload estimation is a useful proxy for direct measures of social awareness. This work validated the assumption that, to accurately estimate another’s workload, a member must understand the other person’s tasks, team relationships, and situation.

According to Kraut et al. (1999), the creation of shared mental models happens through three factors: opportunities to observe, communication, and division of labor within a team. Bolstad & Endsley (1999) found that placing team members physically adjacent allowed more rapid development of shared mental models by creating more opportunities for observation and monitoring.

While physical distance has been explicitly linked to the development of social awareness, social network distance has not. Since social networks are based on the level of communication between members, we hypothesize that shorter social distances between team members will create a higher social awareness. The remainder of this paper documents our efforts to test this theory.

METHOD

This research is the first to use field data sets of temporal organization communications, tasks and training to measure, understand, and describe social awareness. To this point, there has been no experimental (field or laboratory) work, on the relationship between social awareness and communication network evolution. This is due, in part, to the difficulty of large organization network data collection, limited understanding of dynamic network studies, and limited field measures of social awareness. Further, there are simply few suitable databases available to study and describe time-based network evolution and social awareness evolution. As a result, this paper includes a series of field experiments to develop an appropriate dataset. The results from the three experiments were used to understand the relationship between social network distance and social awareness.

The Unit of Action Experiment Series

The three experiments were conducted by the US Army to gain a better understanding of the Unit of Action (UA) Command & Control Structure. The goal is to make decisions about the manning, automation support, and organizational design of the C2 Structure. Normally, the experimental results are used in operations research (ORSA) models, presented in congressional briefings to justify manning and budget requirements, and used to write doctrinal manuals that the Army will use to guide organizational behavior.

Throughout these three experiments, the organizational design under testing was constantly updated. Each experiment resulted in a re-drafted Unit of Action doctrinal manual. Further, each experiment identified tests and conditions that needed to be applied to the next experiment. As a result, both the experiment design and the Unit of Action design were under constant revisement.

The Unit of Action Organization Concept

The Army has, up to now, deployed forces in 2,500 to 4,200-soldier Brigade Combat Teams. These consist of a ground-maneuver brigade (most divisions have three) augmented by other units, such as artillery battalions, which are controlled by the higher commander. The UA is a new “brigade based” structure that will replace the current arrangement, designed for the Cold War when the Army was prepared to fight giant set-piece battles on European soil, and the support roles were organized at the division level.

Brigade combat teams will be restructured into Units of Action. Although the exact configuration of units will vary, the Army has identified a basic Infantry UA design. The Infantry UAs will consist of approximately 3,000 soldiers that include combat, combat support and combat service support functions.

Beyond just new equipment, the UA has far better command and control. Command posts are standardized and integrate enabling capabilities and specialties into command post groupings. Headquarters manning is more robust, experienced, and knowledgeable than in current brigade organizations. Manning is robust enough for 24/7 sustained operations. The staff is more experienced, and enhanced with expertise it did not have -- especially Aviation, PSYOPS, Public Affairs, and Civil Affairs. Attached liaison parties from the Air Force, and from other services and SOF will be more robust. Enhanced battle command networks and functions speed informed decisions, coordination, and execution. Embedded and protected communications nodes insure robust and reliable communications.

The Unit of Action C2 Organization Structure

The experiments reported in this dissertation were not conducted on the entire three thousand soldier UA. Instead, the focus is on the senior leaders of the proposed Unit of Action Command and Control element. That is an element that only makes up approximately fifty of the three-thousand soldiers in the UA. These fifty positions are absolutely critical as they are the few that maintain a top-level view of the organization and interface with the units above and adjacent to the organization. Their primary role is the UA commander’s staff for all operations, intelligence, combat support, personnel, logistics and all combat service support. They are the integrators, coordinators, and planners for everything that happens within the UA.

Each one of the Battle Laboratory experiments had different organizational staff structures under testing. However, they were all fairly similar in concept. Each had a Mobile Command Group(s), Command Integration Cell(s), Maneuver Cell, Logistics Cell, Information Surveillance Reconnaissance Cell, and Effects Cell (table 3.1). While positions and total number of personnel changed in each experiment, these Cell concepts were stable throughout.

UA C2 Element	Description
Mobile Command Group (MCG)	Responsible for the overall mission. The MCG contains the unit commander, deputy commander and a small group of staff to support the commander.
Command Integration Cell (CIC)	Responsible to integrate and double-check all activities of the staff cells. The CIC contains a representative of each of the different functionalities of a Brigade (maneuver, fire support, intelligence, logistics, etc)
Maneuver Cell	Responsible to coordinate the activities of combat forces and combat support forces across the Unit of Action. This cell is oriented on operations.
Information, Surveillance, & Reconnaissance Cell (ISR)	Responsible for intelligence and communication related functions in the UA. Oriented on both facilitating the UA communication and inhibiting the enemy communication.
Sustainment Cell	Responsible for all logistics and combat service support in the UA. Includes supply, maintenance, personnel, and medical.

Table 3.1 Basic Unit of Action cell composition.

A Unit of Action Experiment

To understand and improve the Unit of Action design, the Army is not conducting experiments in the traditional sense. The experimental objective is to create an environment that reflects as closely as possible the conditions expected in battle. There are no experimental & control groups, independent and dependant variables, or treatments.

Instead, the experimenter brings in seasoned military professionals to role-play the required positions. They are given a set of computer based collaboration tools to facilitate coordination (Figure 3.2). A dynamic computer-based simulation is set up to represent blue, red and gray forces on the battlefield. A second set of role-players are brought in to simulate and control the enemy within the simulation based on a broad set of operating parameters and through free-play. Finally, the simulation is started and, periodically, stopped to collect observations from the seasoned military serving as role-players.

Figure 3.2. A typical cell in an Army Battle Laboratory Experiment configured with supporting collaboration equipment.

After the experiment, through the discourse process, seasoned military professionals work out what lessons they learned, which lessons were artifacts of the experiment, and which lessons are relevant to the future UA design. These lessons are then written up and translated into the Army's next doctrine. This process was used for all three experiments. However, each experiment collection was uniquely modified to collect social network data.

Experiment #1 Fort Leavenworth Kansas, October-2003.

On 3 October 2003, 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 3.3). 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.

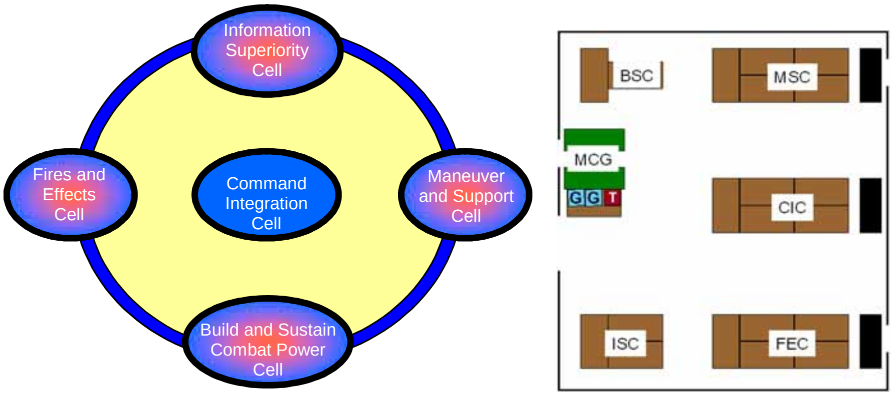


Figure 3.3: Organizational Structure for the BCBL-L exercise. Each Cell consisted of 3-8 agents that could flexibly organize and collaborate as necessary (see table 3.1 for Cell explanations).

Each cell member was not limited by their cell membership and was free to coordinate with other cell members as necessary. The four nodal cells, shown on the outer ring of figure 3, consist of functional groupings of staff roles. The Fire and Effects Cell is responsible for indirect fires and psychological operations. The Maneuver and Support Cell is responsible for ground maneuver and protective activities. The Information Superiority Cell is responsible for the intelligence collection and communication support. The Build and Sustain Combat Power Cell is responsible for logistics. Each of the cell members were expected to identify problems and conduct intra and inter cell collaboration as the battle scenario unfolded.

Data Collection (Experiment #1)

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.

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.

We used a modified version of Entin and Entin's (2001) proximate measure of social awareness in this study. They had found congruence between participants' mutual models and their ability to rate other teammates workload. Operationally, workload congruence 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 3.4). 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.

Workload Questionnaire

Low High
1 2 3 4 5 6 7

Please answer the following questions about WORKLOAD.
Rate Each Question From 1 to 7 (1- lowest, and 7 - highest).

RATE YOURSELF

1 MENTAL DEMAND
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

2 TEMPORAL DEMAND
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

3 EFFORT
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

4 OWN PERFORMANCE
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

5 FRUSTRATION LEVEL
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

6 PHYSICAL DEMAND
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

Figure 3.4: Social Awareness Questionnaire Administered During the Fort Leavenworth Exercise

Social Awareness was calculated 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.

Experiment #2 Fort Lee Virginia, February-2004.

In February 2004, the Ft Lee 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. They were purely interested in the logistics activities associated with the Unit of Action staff. As such this experiment was unique fro the other two in that the majority of the participants had similar logistics background and training. Further all of the tasks injected into the experimental scenario were designed to stress logistics play in the simulations.

Each participant was assigned to a cell with two to five other participants. This experiment was unique from the others in that all people could see all other people, there were no partitions between cells. 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. During their scenario, artificial barriers were in place, 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 (chat, email, etc). A plan-execute-plan-execute cycle was used in the simulation.

There were a total of eight cells in this logistics organization (Figure 3.5). The UA-MCG (Mobile Command Group) consisted of the commander and immediate staff who are responsible for all critical decisions on unit maneuver. Six of the nodes were divided into two teams of three cells. One team focused on aviation logistics (CAB-MCG, CAB-BSC, & AVN-BSC) while the other focused on maneuver logistics (UA-CIC, UA-MSC, & UA-BSC) . The FSB, forward support battalion owned all of the logistics resources necessary for executing logistics support.

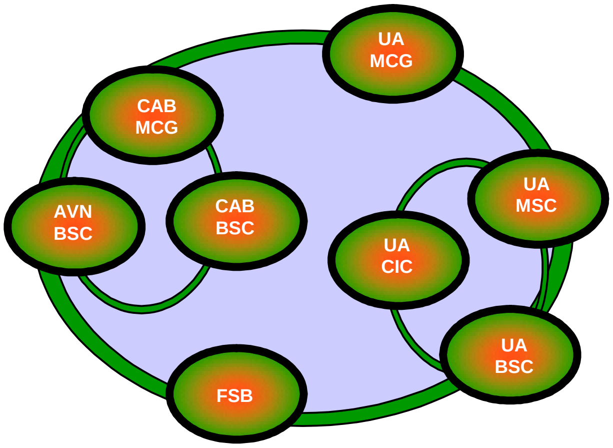


Figure 3.5: Logistical Unit of Action Organization Design.

Data Collection (Experiment #2)

Observations and data collection were conducted over two days immediately following a one-week training period. A plan-execute-plan-execute cycle was used. Data was collected every 60-90 minutes using networked questionnaire that asked the participants for feedback regarding their prior session. Questions included communication frequency, top three tasks worked on, NASA-TLX on self, and NASA-TLX estimation of others. Questionnaire data was collected for a total of 6 sessions of approximately 70 minutes.

Data was collected every 60-90 minutes using networked questionnaire that asked the participants for feedback regarding their prior session. Questions included communication frequency, top three tasks worked on, NASA-TLX on self, and NASA-TLX estimation of others. Questionnaire data was collected for a total of 6 sessions.

Experiment #3 Fort Knox Kentucky, June-2004.

The Unit of Action Maneuver Battle Laboratory (UAMBL) gathered two hundred-twelve experienced army officers to serve as participants in an integrated Unit of Action experiment. Positions ranged from platoon level to division level command and staff groups. The participants were distributed at six locations throughout the United States (figure 3.6) and connected via voice communications, chat, email and shared whiteboard. The cell members operated from mockups of mobile command and control vehicles.

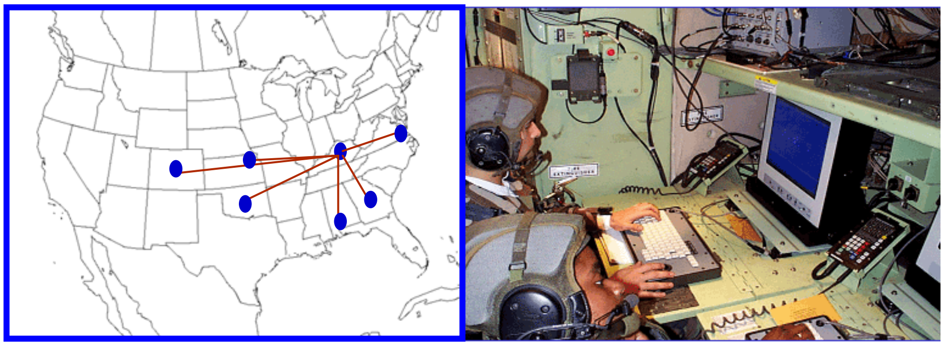


Figure 3.6. The Ft Knox experiment conducted with participants distributed across the United States and positioned in mockups of mobile command and control equipment.

The participants gathered information, coordinated with appropriate staff members, and entered battlefield actions into the simulation. A plan-execute-plan-execute cycle was used with the focus on the UA command and control staff. This was the most robust test of the UA design as there were role-players in multiple subordinate and higher echelon units.

This was the first experiment to test the two command integration cell (CIC) concept (figure 3.7). The two cells were originally designed to provide a 24 hour battle coordination capability. However, during the experiment, the two CIC concept morphed into a planning and execution focus with each cell taking a different responsibility.

.

Figure 3.7. Unit of Action Staff in two CIC configuration.

Data Collection (Experiment #3)

Observations and data collection were conducted every four hours over fifteen days immediately following an eleven day training period. 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.

Questions included communication frequency, top three tasks worked on, top threats to the operation, NASA-TLX on self, and NASA-TLX estimation of their commander. Questionnaire data was collected for a total of 18 sessions. 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.

During this exercise participants only completed workload rating for themselves and their commander using the NASA-TLX. Due to limitations on the collection software, we could not randomize a rating process as in the previous experiments. We, therefore, only measured the organizations social awareness of the organizational commander (leader).

RESULTS

We first studied the correlation between Social Network Distance (SND) and physical distance for all three experiments. We found that the two distance measures have a weak, positive correlation at r = .247 (Expt1); r = .187 (Expt2); r = .295 (Expt3). An extremely high correlation would indicate that people in these experiments were communicating entirely with people they were physically collocated with. However, the weak, positive correlation indicates that while some communication is with the organizational members in close physical proximity, much of the communication is with organizational members outside each person's physical space. This supports the concept that SND and physical distance are not the same.

Next, we tested Olson & Olson (2000) hypothesis that physical distance is the determinant of organizational performance. Recall that social awareness is the proxy for organizational performance and in every experiment we are using same cell membership as a proxy for physical distance. Controlling for the effect of survey period, the effect of physical distance on social awareness accuracy is significant for Expt1at Ft Leavenworth (F [1, 34] =267.47, p < .0001), Expt2 at Ft Lee (F[2,730] = 2.53, p = .08), Expt3 at Ft Knox (F[1, 43] = 3.96, p < .05)(Figure 3.8). Note that in the two experiments where one cell could not observe the other (Ft Leavenworth & Ft Knox), physical distance had a significant effect on social awareness accuracy. In the Ft Lee experiment, where there were no barriers to observation and the participants came from similar backgrounds, the physical distance did not have a significant effect on social awareness accuracy.

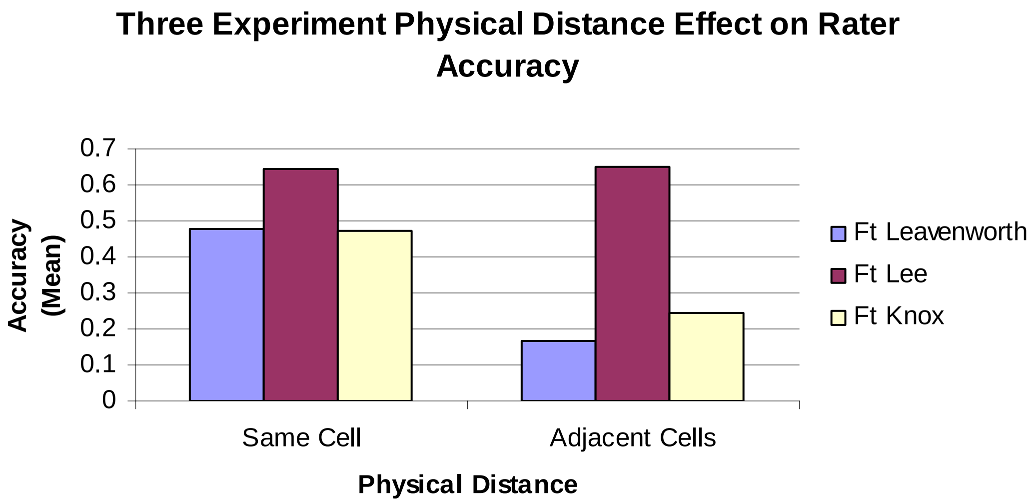


Figure 3.8: Mean social awareness accuracy by physical distance (shared cell vs non-shared cell) (Ft Leavenworth, n = 2210; Ft Lee, n = 715; Ft Knox, n = 213)

Next we tested our hypothesis that social network distance is a predictor of organizational performance. Recall we are using social awareness accuracy as a proxy for organizational performance. Controlling for the effect of survey period, the effect of social network distance is significant for all except one of the experiments: Ft Leavenworth (F[2, 2207] = 180.88, p<.0001, Ft Lee (F [2, 714] = 9.15, p < .0001), Ft Knox (F[2, 208] = .702, p= .058). In every case, social awareness accuracy decreased between people that had more edges in their geodesic (Figure 3.9). However, the change was not significant in the Ft Knox data. Recall that the social awareness measurement for the Ft Knox experiment was different than the other two experiments. The Ft Leavenworth and Ft Lee experiment had the rater evaluate the perceived workload of organizational members randomly selected from the organization. The Ft Knox experiment, however, had the raters always evaluate the perceived workload of the organizational leader.

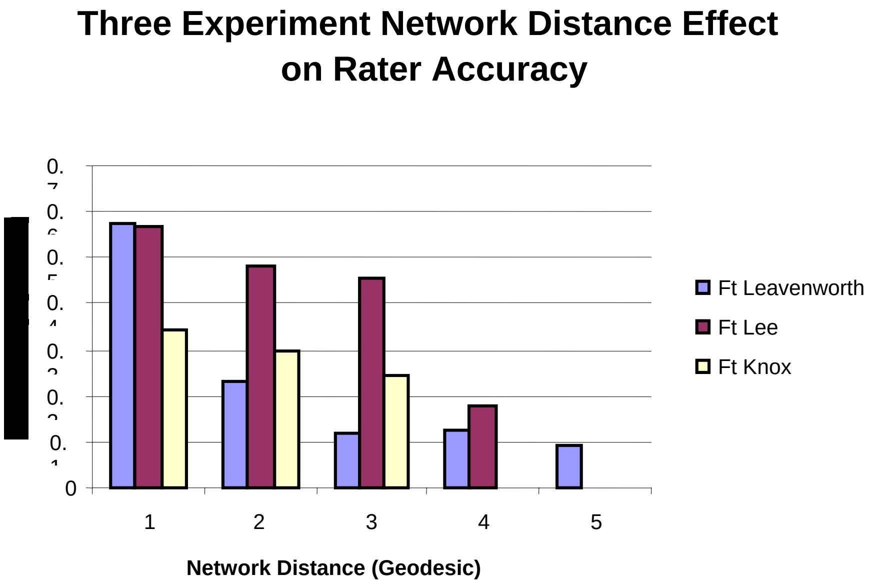


Figure 3.9: Mean social awareness accuracy (Ft Lee n = 715; Ft Leavenworth n = 2210; Ft Knox n = 211)

Also, note that not every experiment allowed for the same social network distance range. Social network distance range is also referred to as the network diameter (Borgatti, 1994; Wasserman & Faust, 1994). The Ft Leavenworth experiment had the largest social network distance range of five, followed by Ft Lee (four), and Ft Knox (three). No single organizational factor determines social network distance range. The social network distance range can vary based on the density of communications in the organization, the number of organizational participants, and the tendency to form subgroups with few people linking the subgroups together.

To account for co-linearity of physical distance & social network distance, account for learning, and account for homophilly we ran a regression model on the predictors of social awareness (Table 3.2). Rater, survey period (session), and homophilly were included to control for the person, the effect of learning, and the effect of background similarity. The results are as in Table 3.2.

	Ft Leavenworth			**Ft Lee**			**Ft Knox**		
Model	r2	n		r2	n		r2	n	
	0.553	1020		0.518	568		0.755	81	
	df	F-Ratio	Beta	df	F-Ratio	Beta	df	F-Ratio	Beta
Network Distance	1	**149.1	0.327	1	**8.83	0.099	1	**7.82	0.461
Physical Distance	1	**136.25	0.286	1	**6.667	0.082	1	0.454	0.2
Homophilly	1	0.427	-0.01572	1	0.001	0.001	1	0.347	0.164
Session	30	**1.86	na	27	**2.69	na	16	**2.362	na

Rater	29	na	na	26	na	na	12	na	na
									****p<.01**

Table 1: Predictors of Mental Model Congruence

Regression results indicate that both physical distance and social network distance are predictors of social awareness, supporting our hypotheses. Under this model, social network distance is a significant ($p < .01$) predictor of social awareness for all experiments, and physical distance is a significant predictor ($p < .01$) of social awareness in all experiments except Ft Knox. The mean congruence of individuals in the same cell is 10.65, which is significantly more congruent than those in different cells (mean = 11.61).

DISCUSSION

Both social network distance and physical distance were predictors of social awareness accuracy. However these distance measures are qualitatively different from each other. Physical distance provides information about who is likely to use face-to-face coordination and who is proximate. Social network distance is capable of providing information about who is linked by the task structure, context, and organizational structure. The weak correlation between the two distance measures further supports the independence of social network distance and physical distance as predictors of social awareness.

Physical Distance or Observability

Experiment #2- Ft Lee is the one exception to the physical distance matters premise. Recall that the experiment #2 design differed in that although divided into physically seperated cells, there were no dividers between the cells. As a result, it was easy for most organizational members to observe the other organizational members. It appears that observability, even from a distance, offers enough information to support social awareness accuracy. Therefore, we have some evidence that, observability, not physical distance, matters in the support of organizational performance.

Past experiments on small teams have demonstrated the importance of observability. What is not clear is whether observing others behaviors or observing others workspace is critical to performance. Bolstad & Endsley (1999) found enabling team members in observe one another’s actions improved team performance. Kraut et al (1999) found that allowing team members to observe team workspace also improved team performance. What is clear is that observeability, which can be a characteristic of physical distance, is a key mechanism in social awareness.

Social Network Distance to Key Leadership

Experiment #1 & Experiment #2 demonstrated a significant decrease in social awareness with increased geodesic distance. However, Experiment #3-Ft Knox did not have a significant decrease in social awareness with increased geodesic. Recall that the experiment #3 design differed in that we only measured the organizations social awareness of the leader. It appears that there are qualities of the leader –

subordinate relationship that support the maintenance of social awareness of the leader over longer social network distances.

In our own observations of the command and control experiments, we found that leaders tend to spend a great deal of time articulating their point of view and stating their current work priorities. In our pilot study we found that leaders tend to initiate many small communications to many members of the organization (Graham et al, 2002). As a result, while a organizational member may be a far network distance from the leader, they may also be able to observe other organizational members communicating with the leader. This second person observation may supply enough social awareness information of the leader to decrease the significance of social network distance.

Homophilly and Social Awareness

None of the experiments demonstrated a significant relationship between background similarity/ homophilly and social awareness. Recall that homophilly refers to the similarity (demographics, training, experiences, etc) between any two members of the organzation. For these experiments similarity of training specialty/branch (infantry, armor, logisitics, etc), was used to determine homophilly.

However, there is evidence that homophilly did determine who each organizational member chose to communicate with (Figure 3.9). For all experiments the highest average homophilly was between members that were in direct communication (social network distance = 1). There was a significant decrease (p < .01) in mean homophilly as the network distance increased across all experiments.

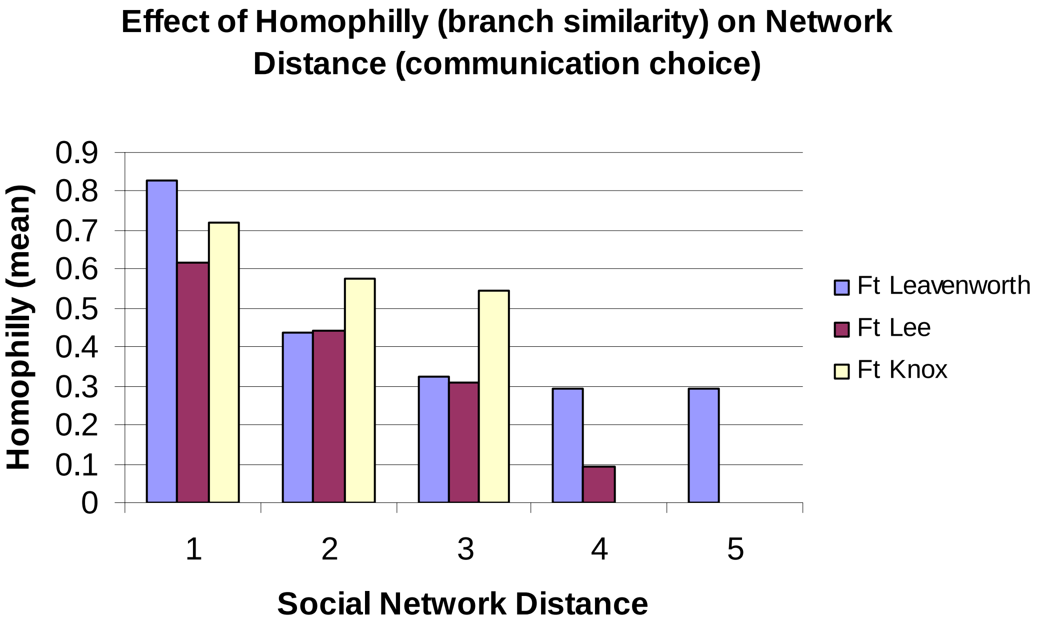


Figure 3.9. Effect of Homophilly on choice of communication partner and within each social network distance.

This finding may be an artifact of the experiemental context. While toplevel command and control decisions are made by integrating the input of multiple training branches (infantry, armor, logistics, etc), most of the low-level problem solving is between members of the same training branch. However, previous studies in other contexts have found homophilly as a significant predictor of communication choice in networks (Monge & Contractor, 2001). People are drawn to communicate with others they find similar to themselves. As a result, homophilly is a strong determinant of who directly communicates, and therefore, has improved social awareness.

Social Awareness & Learning

There was a significant improvement in social awareness over time (table 1 & Figure x). The improvment indicates that there was a learning effect over the course of the experiment. Through the normal course of interaction and problem-solving, the participants had first and second hand access to the quirks, body & voice indicators, and task requirements of the other organizational members.

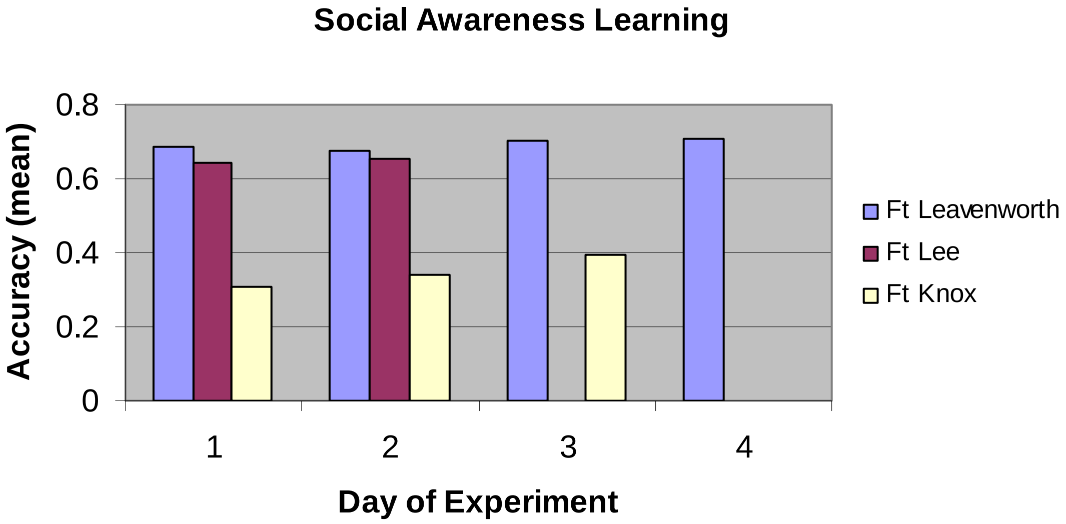


Figure x. Social Awareness Learning trend over three experiments. Note: there is a restart cost on each day so that the first collection of the day is lower than the others.

Clark & Brennan (2000) found that as people spent more time together they learned more about one another. This knowledge improves the grounding and therefore performance of small teams. Our findings also indicate that time is a factor for social awareness improvement across a large organization. With more time, each member has more accurate social awareness of members of all members of the organization, whether or not they have direct contact.

Limitations

This research was based on measurements of three organizations over a relatively short period of time. In every experiment, the organizational members were assigned roles and introduced to one another only a few days before the experiment began. The fact that social awareness accuracy significantly improved over time indicates that the organization had not stablilized. Therefore the results may not completely translate to organizations that have operated together for extended periods of time.

Also, the use of self reports is not an optimal method to generate social network graphs. Studies on social network analysis collection (Obradovich, et al, 2004) found that self reports suffer from the primacy recency effect. Under primacy recency, individuals tend to accuratley recall the first and last activities in a session, but less accurately recall the activities in between. As a result, our social network distance measure may be skewed by the participants recall of their own communications during each session. Direct collection of communication through voice transcripts, email logs, and chat logs would provide a more reliable data source.

Last, these experiments are only slightly removed from collection in the field. Similar to a real organization out of the laboratory, there are few controls and constants in the experiment. Participants have varying degrees of expertise in their designated training branch, may have prior relationships from previous work assignments, and have varying degrees of comfort with communication technologies. The experimental scenario is stocastically evolving as decisions are entered into the computer and the automated enemy forces react. This causes a great deal of variability in the tasks the the organizational members identify and problems they choose to solve.

CONCLUSION

Our results did show a significant increase in social awareness as both social network distance decreased and physical distance decreased. Apparently, both constructs play a part in social awareness accuracy. Thus, we should consider the qualititative aspects of physical distance and social network distance.

Physical distance is essentially immutable and static. A person’s location is often driven by the best position to collect information, not collaborate about the information. Therefore, it is unlikely that key people can simply meet face-to-face in the same room every time their mental model congruence falls below some threshold. However, given the finding that observability is a key characteristic of physical distance, we have other options. We may be able to increase the observability of others and their workspace through collaborative tool design. By providing a view into the workspace of others, we may increase social awareness accuracy and team performance.

Unlike physical distance, social network distance, as a function of task, context, and organizational structure, is dynamic and is under greater personal control. However, our findings indicate that people are more likely to communicate with similar organizational members. By providing forcing functions to spread direct communication, social network distance will decrease between distant organizational members. The tradeoff is that with increased communication, the workload of the agents increases as well. As we move to network organizations (Nohira & Eccles, 1992, Miles & Snow, 1995), SND will become more critical as a design tool for organizational structure and collaborative tool selection. Finding ways to shorten the social network distance without increasing workload, and finding ways to increase mental model congruence without high training costs, are worthwhile areas for future research.

There is an old saying, “it is not what you know, it is who you know.” We would add a caveat: ‘it is also who knows those you know’. This research suggests that social network distance may be more important than physical distance for understanding group performance and organizational structure. We found that physical distance and social network distance are independent from each other, yet they are both good predictors of social awareness. However, social network distance has inherent qualities that allow for controls and support to improving social awareness.

ACKNOWLEDGEMENTS

We would like to thank Diane Ungvarsky the Battle Command Battle Lab.

REFERENCES

Bolstad, C. A., & Endsley, M. R. (1999) Shared mental models and shared displays: An empirical evaluation of team performance. In *Proceedings of the 43rd Annual Meeting of the Human Factors and Ergonomics Society.* Santa Monica, CA.

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

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

Dorish P., & Bellotti, V. (1992). Awareness and co-ordination in shared workspaces. In *Proceedings from CSCW'92*, Toronto. ACM press

Entin, E.E. (1999). Optimized command and control architectures for improved process and performance. In *Proceedings of the 1999 Command and Control Research and Technology Symposium*. United States Naval War College, Newport, RI.

Entin, E.E and Serfaty, D. (1999) Adaptive team coordination. *Human Factors*, 41, 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.

Gutwin C., Greenberg, S. & Roseman, M. (1996). Workspace Awareness in Real-Time Distributed Groupware: Framework, Widgets, and Evaluation. In *proceedings from HCI'96,* London, UK.

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*, 139-183. Elsevier Science: NorthHolland.

Krackhardt, D. (1994) Graph Theoretical Dimensions of Informal Organizations. In *Computational Organization Theory*. K.M. Carley and M.J. Pretula, eds., 89-111. Hillsdale, NJ: Lawrence Erlbaum Assoc.

Kraut, R. E., Fish, R.S., Root, R.W. & Chalfonte, B.L. (1990) Informal Communication in Organizations: Form, Function, and Technology. *In S Oskamp and S Spacapan (eds): People's Reactions to Technology*. Newbury Park, CA: Sage Publications.

Kraut, R. E., Fussell, S. R., & Lerch, F. J. (1999) Shared mental models in work groups: A controlled field study. *The Symposium on Shared Cognition in Teams: Predictors, Processes, and Consequences at the Annual Meeting of the Society for Industrial and Organizational Psychology (SIOP),* Atlanta, GA

Miles, R. E. & Snow, C. C.. (1992) Causes of failure in network organizations, *California Management Review*, vol. 34, pp. 53-71.

Monge, P. R., & Contractor, N. S. (2001). *Emergence of communication networks.* In F. M. Jablin & L. L. Putnam (Eds.), The New Handbook of Organizational Communication: Advances in Theory, Research, and Methods (2nd ed., pp. 440- 502). Thousand Oaks, CA: Sage.

Moran, T.P. & Andersson, R.J. (1990). The workday as a paradigm for CSCW. In *Proceedings from CSCW'90*, LA, California, ACM Press.

Nohria N, Eccles RG. (1992) *Networks and Organizations: Structure, Form and Action*. Boston, Mass.: Harvard Business School Press.

Olson, G. M., & Olson, J. S. (2000) Distance Matters. *Human-Computer Interaction. Special Issue: New agendas for human-computer interaction*, 15(2-3), 139-178.

Scott, J., (1992) *Social Network Analysis*. Newbury Park CA: Sage.

TOLLMAR K, SANDOR O and SCHÖMER A (1996). Supporting social awareness @work – design and experience. In *Proceedings from CSCW'96,* 298-307, Cambridge, MA, USA. ACM press.

Wasserman, S. and Faust, K. (1994) *Social Network Analysis: Methods and Applications.* Cambridge: Cambridge University Press.

Dynamic Network Analysis of the Network-Centric Organization: Towards an Understanding of Cognition & Performance

Chapter 5

Towards Field Application: Dynamic Network Theory, Naturalistic Decision Making, and Organizational Analysis in Network Organizations

Abstract

Naturalistic decision making can be characterized as focused on expertise and the ‘what you know’ while dynamic network theory is descriptive of the organization and the ‘who you know’ (Carley, 2003). Both theories share a common thread in that they seek to understand behavior and decision-making as it really happens in the world as opposed to a laboratory-based normative model. Both approaches are challenged by the transition of our military into network-centric organizations. Network organizations, by eliminating traditional strong hierarchy and management stovepipes, allow members at the lowest levels to collaborate and solve problems (Nohria & Eccles, 1992; Podolny & Page, 1998). Dynamic network analysis allows researchers to simultaneously understand individual and social characteristics relevant to network organizational behavior. This paper describes the application and integration of traditional naturalistic decision making methods and dynamic network analysis to the study of decision-making in prototype network organization experiments. Beyond a single experiment, we compare two experimental command and control organizations as tested during US Army Battle Laboratory experiments. The analysis shown here was applied to design decisions about a prototype network organizational system in the Army.

I. Introduction

NDM would benefit from a methodology that effectively describes a large organization as it evolves during the execution of a task. Naturalistic Decision Making (NDM) research has largely based on observations of individuals or small teams in the field. Unfortunately, in large scale and network oriented organizations, observation may not be sufficient. Ethnographers can only be in one location at a time so that their observations are restricted to the part of the organization they are observing. Researchers have sought to overcome this problem by 1) using synchronized video/audio tapes, 2) employing multiple observers, 3) periodically moving the observation point of a observer, and 4) conducting post-hoc interviews. In a large organization, the first two methods are resource intensive. For instance, a recent field observation of the Recognition Primed Planning Process in a 56 member organization required 15 observers simultaneously inputting to 8 computer logs over the course of four days (Ross et al, 2004). It also took a large number of post-experiment data analyst resources to put together a coherent picture of the data and develop conclusions. The third and fourth methods suffer from different collection problems. In periodically moving the observation position, the observer has to assume that the organization form is static and predetermined as they move from point to point in the organization. In the case of post-hoc interviews, it has been found that participant's recall is strongly affected by the primacy-recency effect (Obradovich, Schneider, Graham, & Gonzalez, 2004). Thus the results from post-hoc interviews would not accurately reflect the activities in the organization during the execution of a task.

This paper introduces the dynamic network theory and its use in large, dynamic organizations. Through the description of two important organizational experiments in military decision making we will demonstrate that dynamic network theory supports current NDM methods by providing an overall view of a large organization's interactions and decisions. This paper will first describe two prototype network organizations tested by the US Army. We then describe application of three dynamic network analyses: temporal network density, network-centric organizational behavior, and key temporal problem-solving groups and individuals. Lastly, we will discuss how dynamic network theory augments traditional naturalistic decision-making data collection and analysis methods.

II. Dynamic Network Theory

Dynamic network analysis holds particular promise for understanding and describing the newest organizational form: the network organization. Network organizations are unconstrained by traditional organizational charts and hierarchies in deference to self-organizing problem-solving groups (Podolny & Page, 1998). Dynamic network theory includes the best of social psychology-based network theory integrated with the most applicable cognitive psychology (Carley & Remiga, 2004). In applying dynamic network analysis, a researcher is able to simultaneously understand characteristics of the individual agent and the organization the agent resides within.

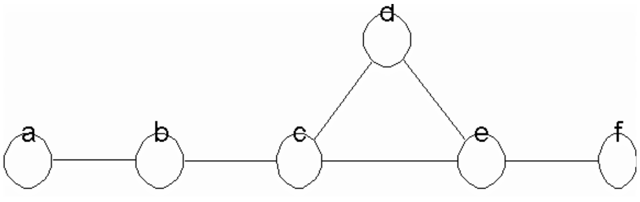


Figure 1. A simple network graph (Borgatti, 1994)

A social network is a graph consisting of individuals and connections between them(Scott, 1992). In a social network graph, individuals are represented as nodes and communication between individuals is represented by links between the nodes (Figure 1) (Borgatti, 1994; Scott 1992). Communication data can be gathered by looking at various measures such as shared email headers or phone calls, or by surveying the members (Wasserman & Faust, 1994).

Dynamic Network Theory has provided a set of social as well as cognitive measures. Among the most important social measures is the social network distance. This measure also referred to as the geodesic, is the number of links or actors between two members of a social network graph (Borgatti, 1994). Geodesics are critical in calculating member centrality; how important a member is to the network (Freeman, 1979) and information diffusion; how well information moves through an organization (Sprang & Tuma, 1993).

We have used network graphs to graphically represent the dynamic relationships occurring between organizational members in large, dynamic organizations(Graham, Gonzalez & Doyle, 2003; Graham, Schneider, Bauer, Bessiere & Gonzalez, 2004). Individual cognitive measures can also be inferred from the network graphs. Using simple network graphs augmented with knowledge and task structures it is possible to produce cognitive load estimates for an individual (Carley & Remiga, 2004) or estimate transactive memory of individuals in an organization(Liang, Moreland & Argote, 1995). Therefore, dynamic network analysis can simultaneously describe individual and organizational characteristics.

III. Collecting Social Network Data: Two Prototype Command & Control Organizations

The US Army is in the infancy of a ten-year organizational design process for a new command and control element. The concept, dubbed the Unit of Action for the Future Force, allows the organization to perform in a network-centric fashion (Podloni & Page, 1994). A network organization can flexibly organize to solve problems at the lowest level without the direct supervision of a manager or the constraints of typical stovepipes. This new organizational concept is a severe departure from the traditional military hierarchy.

The US Army tests new organization designs by running role-player driven simulations. Experts from throughout the Army are assembled at a Battle Laboratory and assigned to positions in the prototype organization. Our team participated as data collectors and analysts for multiple US Army prototype organization experiments. We present here the data collected during two such experiments on a prototype organization called the Unit of Action staff. The Unit of Action staff is under study as a replacement to the standard Army brigade staff hierarchy.

During the short experimental life of the organization, we collect social network data by asking participants to report their communications during a specific period. The responses are ordered by the frequency of communication during the previous session. In addition, workload data was gathered using the NASA TLX (Task Load Index) (Hart & Staveland, 1988) assessment consisting of seven workload parameters on a Likert scale.

Participants were asked to rate their own workload as well as that of five other randomly selected participants. The ratings of others' workload were used to estimate organizational state awareness and will be discussed in the last section of this chapter. We also collected the number and type of tasks performed by each individual. Observers were also posted throughout the organization to collect and report on the participants activities and conduct interviews as time permitted.

The two organizations were both considered the leadership staff of Units of Action but with a critical difference. One consisted of all the staff elements within an Army Unit of Action; operations, planning, logistics, intelligence, maneuver, etc. While the other consisted of only the logistical elements within the Unit of Action. The first was prototyped and tested at the Ft Leavenworth Battle Command Battle Laboratory (BCBL) and the second was prototyped at the Ft Lee Combat Service Support Battle Laboratory (CSSBL).

First Organizational Prototype: Battle Command Battle Laboratory (BCBL)-Leavenworth

The Fort Leavenworth Battle Command Battle Laboratory (BCBL) gathered fifty-six experienced Army officers to serve as staff members for a Unit of Action command and control staff. Each participant was assigned to one of five functional cells with three to eight other participants. The participants gathered information, coordinated with appropriate staff members, and made decisions as the computer based battle scenario required. 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 (radio and chat through the computer).

The hub of the organization was the command integration cell (see Figure 2). The command integration cell served as the final approval on all decisions and members were expected to monitor all problem-solving traffic between the nodal cells. The command integration cell was unique in that it had representatives of each of the nodal cells.

Figure 2: The Command Integration cell in the Command post exercise

The four nodal cells, shown on the outer ring of Figure 3, consist of functional groupings of staff roles. The Fire and Effects Cell is responsible for indirect fires and psychological operations. The Maneuver and Support Cell is responsible for ground maneuver and protective activities. The Information Superiority Cell is responsible for the intelligence collection and communication support. The Build and Sustain Combat Power Cell is responsible for logistics. Each of the cell members were expected to identify problems and conduct intra and inter cell collaboration as the battle scenario unfolded.

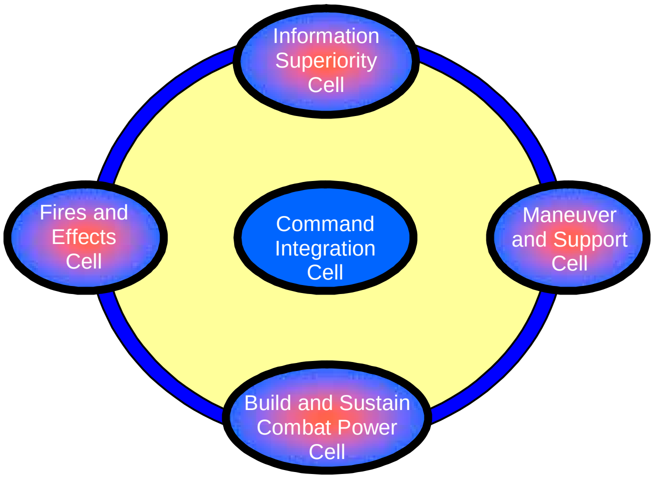


Figure 3: Organizational Structure for the BCBL-L exercise. Each Cell consisted of 3-8 agents that could flexibly organize and collaborate as necessary.

The observations and data collection were conducted over four days immediately following the one-week training period. A plan-execute-plan-execute cycle was used.

Data was collected every 60-90 minutes using networked questionnaire that asked the participants for feedback regarding the prior session. Questions included communication frequency, NASA-TLX self-rating, and NASA-TLX other estimation. Questionnaire data was collected for a total of 16 sessions.

Second Organizational Prototype: Combat Service Support Battle Laboratory (CSSBL)

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. During their scenario, artificial barriers were in place, 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 (chat, email, etc).

There were a total of eight cells in this logistics organization (Figure 4). The UA-MCG (Mobile Command Group) consisted of the commander and immediate staff who are responsible for all critical decisions on unit maneuver. Six of the nodes were divided into two teams of three cells. One team focused on aviation logistics (CAB-MCG, CAB-BSC, & AVN-BSC) while the other focused on maneuver logistics (UA-CIC, UA-MSC, & UA-BSC). The FSB, forward support battalion owned all of the logistics resources necessary for executing logistics support. Each cell member was not limited by their cell membership and was free to coordinate with other cell members as necessary.

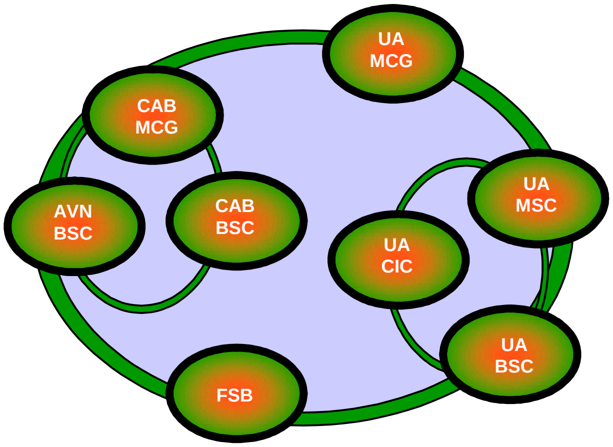


Figure 4: Logistical Unit of Action Organization Design.

Observations and data collection were conducted over two days immediately following a one-week training period. A plan-execute-plan-execute cycle was used. Data was collected every 60-90 minutes using networked questionnaire that asked the participants for feedback regarding their prior session. Questions included communication frequency, top three tasks worked on, NASA-TLX on self, and NASA-TLX estimation of others. Questionnaire data was collected for a total of 6 sessions of approximately 70 minutes.

III. Using Dynamic network analysis

Applications of Dynamic Network Analysis

Using Social Network Analysis, we are able to make conclusions about the organizations under study which are not available via NDM. In this section we show how SNA reveals effects of homogeneity on organizational learning, how we locate unexpected leaders in the organization, and how we identify critical information feeders who might otherwise go unnoticed.

Explaining Organizational Learning using Network Density

In using nearly the same collection and analysis techniques across the two organizations, we demonstrate differences in organization learning. The specific dynamic network measure used in this analysis is network density (Freeman, 1979). Network density is the sum of active links in an organization divided by all potential links between members. A fully dense network/organization would have every person (node) linked to every other person.

In a newly formed organization, behavioral norms have not yet been established (Moreland, 1999). For instance people do not know who has expertise on different topics, or who they can or should contact when attempting to solve a specific problem. As the organization gains experience, we expect to see meaningful changes in organizational behavior as measured by network density.

Recall that the participants in these studies are operating as an organization for the first time. We expect a new organization to start out with a relatively sparse network, gradually increase linkages as the members explore and establish necessary communication channels, and then peak at some level of density. Once the organizational norms are established, we expect all variations in behavior are attributable to factors such as organizational task load.

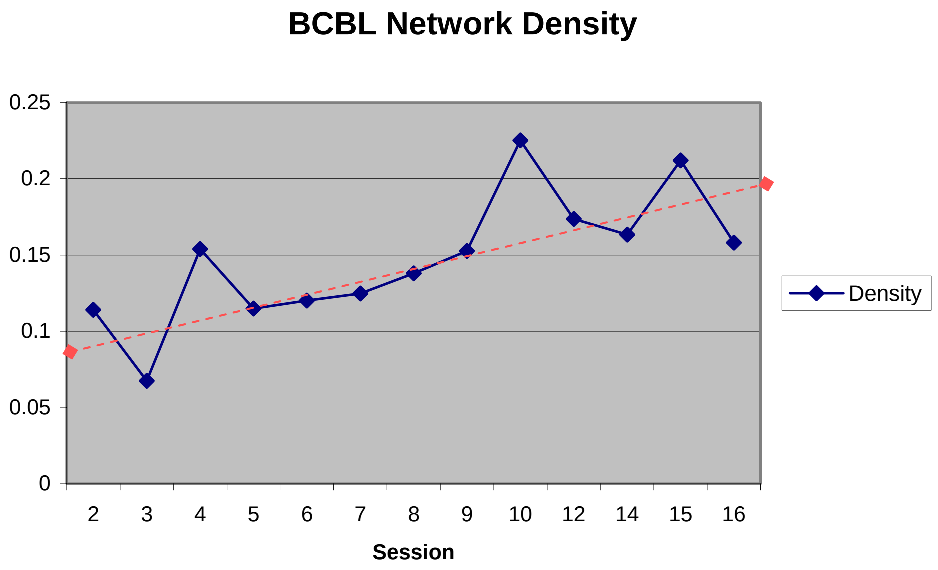


Figure 5. Network Density change over the course of the BCBL-Leavenworth experiment.

Figure 5. shows the change in network density in our first data set over the course of a four day scenario. Network density starts at approximately 10% and finishes at approximately 20%. The communication density peaks at time period ten. Observer reports indicate that the variations starting at time period ten can be attributed to changes in the scenario task load. Variations prior to time period ten are more likely an outcome of learning and the creation of organizational norms. A critical point to take from this data is that the final sessions, after the norms are established, are potentially closer to how this organization would operate than the beginning sessions.

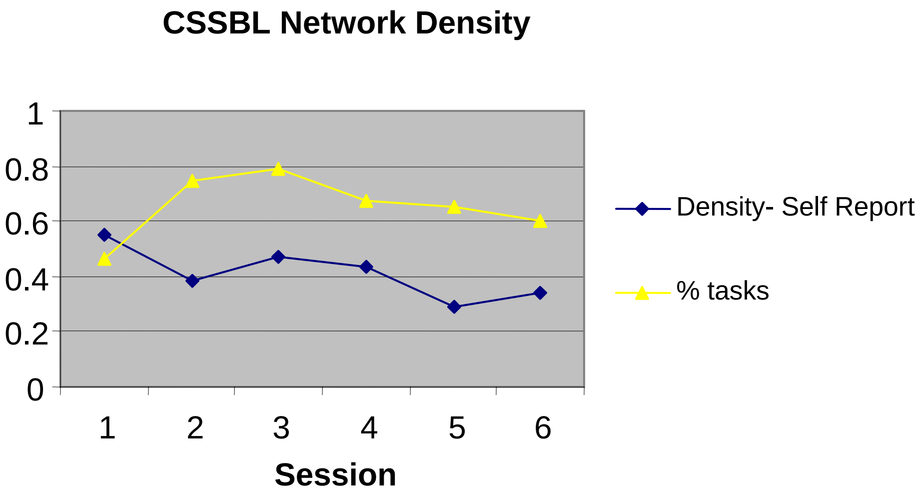


Figure 6. Task- Communication Network Density Change over the Six MSL inject sessions. Note that the change in number of tasks per session is reflected in the communication density

In the CSSCL-Lee scenario, we found a different pattern of organizational learning. Figure 6 shows the change in network density over the six sessions of the experiment. No clear pattern emerges in network density with respect to time (i.e. learning). However the network density does follow the changes in task load for all sessions except the first. It appears this organization established organizational norms more quickly that the BCBL scenario.

We explain the major difference between the BCBL-Leavenworth and CSSCL-Lee organizations patterns of organizational learning by noting the degree to which the prototype organizations are analogous to the legacy force structure. The CSSCL-Lee organization and positions are very similar to the legacy force. Therefore participants in that experiment can transfer their existing knowledge of the roles, tasks, and organizational norms. However the structure used in the BCBL-Leavenworth experiment is a radical departure from the legacy force and required the participants to learn new information before organizational norms could be established clearly.

Locating Hidden Organizational Structure

As an organizational ethnographer cannot be everywhere at once, it is important to understand the relationship between all agents and the ones under observation. Traditionally, we can use an organizational diagram or question the members of the organization. However, these methods do not provide the true structure of the organization. Krackhart (1987) found that experienced managers were only 80% correct in identifying the formal and informal leaders in their organization. Furthermore, these positions and relationships are constantly changing in a dynamic decision-making environment such as the future force structure.

We can use communication networks to get at actual dynamic organizational structure by computing various network measures on the organization. This section will first show the application of triads to the BCBL-Leavenworth data set and then demonstrate the application of betweenness centrality to the CSSBL-Lee data set.

To better understand the network structure we look at triads. A triad is a group of three agents tightly bonded through reciprocal communications. In this futuristic context, problem-solving by a single individual is the domain of expert systems and artificial intelligence. As a result, in the prototype organizations, triads of participants are important as complex decisions are made through collaboration and teamwork. In the complete UA data, we looked for groups or triads that had the highest frequency of occurrence during the life of the organization.

In a stable and hierarchical organization, we would expect most triads to persist over time. In the Ft Leavenworth data, 302 unique triads formed over the course of the experiment. Of those 209 (69.2%) were unique to a particular session. This high uniqueness indicates a flexible organization reacting to the changing requirements of the system.

In the Ft Leavenworth data, one triad existed across eight of thirteen observed organization sessions. This triad consisted of the CIC-Executive Officer, MSC-Operations Officer, and the Sustainment Officer (Figure 7). No other triad occurred in more than 1/3 of the sessions.

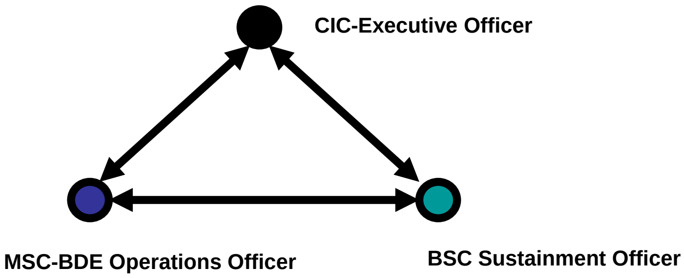


Figure 7. UA primary command group information hub as found during experimentation by the BCBL-Leavenworth.

Here is where the combination of observation and dynamic network analysis proved invaluable. This triad was not evident in the observational data. This data was missed because the each of the triad members are in geographically separated cells. Further, in the legacy staff structure, the Sustainment Officer would not be considered a likely member of the most dominant decision-making team, thus creating an experimenter bias in the observers. Once the finding was brought to the observers, interviews were conducted that determined that the Executive Officer and Operations

Officer worked tightly with the Sustainment Officer to understand the logistics implications of their decisions.

Locating the otherwise invisible information nodes

While triads can give insight to the organization structure and problem-solving teams, another measure called betweenness centrality gives insight on specific actors. Betweenness centrality the degree to which one actor falls on the shortest paths between other pairs of actors in the network (Freeman, 1979). That is, the more people depend on me to make connections with other people, the more power I have. In a command and control organization, people with high betweenness can be the integrators of information and decisions while people with low betweenness can be the feeders of information and expertise. The following sections give examples of using betweenness centrality measures and observations to understand these complex relationships.

Identifying Unexpected Leaders

In addition to the triad analysis, betweenness centrality provides another method for identifying unexpected leaders.   Table 1 shows the top four of fifty-six participants in the BCBL-Leavenworth experiment ranked by betweenness centrality.

Betweenness Centrality Ranking
1 - Operations Officer
2 - Executive Officer
3-   Space Officer
4-   Intelligence Officer

Table 1: Top four of fifty-six participants in the BCBL-Leavenworth experiment ranked by betweenness centrality

In both legacy force and future force scenarios, it is normally expected that Commander, Operations Officer, Executive Officer, and Intelligence Officer would be the leaders of the organization. In Table 1, three of the four expected leaders appear (Operations Officer, Executive Officer, and Intelligence Officer). However, the Commander is missing and the Space Officer is present.

The role of the Space Officer was originally designed as a purely functional and specialized role which did not include organizational leadership. However in this experiment his expertise was instrumental in multiple important tasks. He had access to resources which supported planning, execution, and assessment of battlefield operations that proved critical to operational success. As a result he unexpectedly became a de facto integrator and leader within the organization.

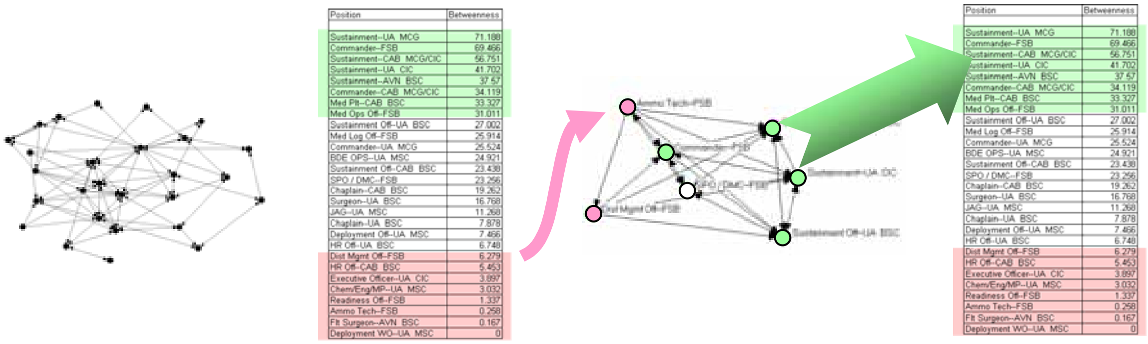
The unexpected disappearance of the Commander from table 1 is explained by the characteristics of the betweenness centrality measurement and by the characteristics of military command and control. Betweenness centrality is essentially a measure of information brokerage, which is moving information from one person to another. As such we would not expect a commander of a military organization to be among the highest in betweenness centrality, as a commander's role is to formulate a plan and monitor its execution by his subordinates, but the coordination of particular tasks is left to his subordinate leaders such as the Executive Officer. Therefore he does not act as an information broker.

A traditional NDM analysis based on observation would not be expected to identify the Space Officer as a leader, because without close observation of the Space

Officer this function would have remained hidden. As well, betweenness centrality offers an insight into the nature of the Commander's role that might not have been otherwise noticed.

Identifying Information Feeders

Often an organizational role can be considered unimportant due to observed behaviors and low betweenness ranking. This was the case for the ammunition technician in the logistics Ft Leavenworth organization. The ammunition technician provides logistics coordination and expertise for all matters related to ammunition. Observation data on the ammunition technician indicated that the position was peripheral and provided little contribution to the organization. In fact, observers were unwilling to use very much of their time on the ammunition technician position. Our own analysis indicated that the ammunition technician was third from the bottom in betweenness (Figure 8). Combined with the observation data, the indication was that the ammunition technician did not contribute significantly to staff decision-making.



Position	Betweenness
Sustainment--UA MCG	71.188
Commander--FSB	69.466
Sustainment--CAB MCG/CIC	56.751
Sustainment--UA CIC	41.702
Sustainment--AVN BSC	37.57
Commander--CAB MCG/CIC	34.119
Med Plt--CAB BSC	33.327
Med Ops Off--FSB	31.011
Sustainment Off--UA BSC	27.002
Med Log Off--FSB	25.914
Commander--UA MCG	25.524
BDE OPS--UA MSC	24.921
Sustainment Off--CAB BSC	23.438
SPO / DMC--FSB	23.256
Chaplain--CAB BSC	19.262
Surgeon--UA BSC	16.768
JAG--UA MSC	11.268
Chaplain--UA BSC	7.878
Deployment Off--UA MSC	7.466
HR Off--UA BSC	6.748
Dist Mgmt Off--FSB	6.279
HR Off--CAB BSC	5.453
Executive Officer--UA CIC	3.897
Chem/Eng/MP--UA MSC	3.032
Readiness Off--FSB	1.337
Ammo Tech--FSB	0.258
Flt Surgeon--AVN BSC	0.167
Deployment WO--UA MSC	0

Position	Betweenness
Sustainment--UA MCG	71.188
Commander--FSB	69.466
Sustainment--CAB MCG/CIC	56.751
Sustainment--UA CIC	41.702
Sustainment--AVN BSC	37.57
Commander--CAB MCG/CIC	34.119
Med Plt--CAB BSC	33.327
Med Ops Off--FSB	31.011
Sustainment Off--UA BSC	27.002
Med Log Off--FSB	25.914
Commander--UA MCG	25.524
BDE OPS--UA MSC	24.921
Sustainment Off--CAB BSC	23.438
SPO / DMC--FSB	23.256
Chaplain--CAB BSC	19.262
Surgeon--UA BSC	16.768
JAG--UA MSC	11.268
Chaplain--UA BSC	7.878
Deployment Off--UA MSC	7.466
HR Off--UA BSC	6.748
Dist Mgmt Off--FSB	6.279
HR Off--CAB BSC	5.453
Executive Officer--UA CIC	3.897
Chem/Eng/MP--UA MSC	3.032
Readiness Off--FSB	1.337
Ammo Tech--FSB	0.258
Flt Surgeon--AVN BSC	0.167
Deployment WO--UA MSC	0

a. b. c. d.

Figure 8. From the total 28 person communication network (a), we can construct a ranking of the betweeneness centrality of each member (b). People in the light green are

high in betweeness and are considered integrators of information. Agents in the pink are low in betweeness and are either feeders of information or simply not important to the organizations objectives. The third graph (c) is the ego network of the ammunition technician. While he is low in betweenness, fifty percent of his ego network consists of members in the top of the organization (d).

To test this hypothesis, we construct an ego network of the ammunition technician (c). An ego network consists solely of the agents that a single agent communicates with (Wasserman, 1994). Surprisingly, the ego network data is in direct contrast to the initial observations of the ammunition technician position. Fifty percent of his ego network is with agents that are in the top of the organization. So while the ammunition officer is certainly not the hub of the organization, his input is valued by four of the top six decision-makers of the organization.

Neither observation nor surface features of the communication network make this finding apparent. Post-interviews with the ammunition technician did support the critical nature of the position, but then again, every agent interviewed indicated their position was equally critical. Only by going deeper into a qualitative analysis of the communication network was our initial hypothesis found unsupported.

V. Discussion and Conclusion

All of the methods applied here were generated post-hoc using self-report questionnaires. While the results are valuable, additional gains can be attained by making the collection and analysis process as real-time as possible. Further, it is important to eliminate periodic questionnaires that disrupt the normal organizational activities.

The proliferation of distributed network organizations has resulted in supporting electronic collaborative tools. We are currently experimenting with using the real-time electronic communications record to conduct dynamic network data collection and analysis. One goal is to eliminate the invasiveness and disruption of a questionnaire. Another goal is to provide the organizational observers with visibility over the organization network and participants so they can better study emergent behavior not readily apparent in the observations.

There is an old colloquium, ‘its not what you know, but who you know’. In network organizations, the truth is that ‘who’ and ‘what’ vary in importance based on the dynamic context. NDM has been a boon to explaining what decision-makers know and how they apply that knowledge. Combined with Dynamic Network Analysis, we can explain how ‘what you know’ integrates with ‘who you know’. In fact, describing decision-making, separate from the in situ organizational state provides for flawed understanding.

The US Military is just beginning to experiment with the Recognition-Primed Planning Process at the same time it is converting to the network organization concept. Integrating NDM with Dynamic network theory provides the tools and methodology for understanding the implications at the individual and organizational levels. Understanding the macro-cognitive in concert with the individual cognitive can only lead to better systems.

REFERENCES

Bolstad, C., Schneider, M., Graham, J., & Gonzalez, C. (2004) Collaborative Tools & Shared Mental Models. In Proceedings of the 2004 Human Factors & Ergonomics Society, New Orleans, LA.

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

Carley, Kathleen M. and Jeff Reminga, (2004). ORA: Organization Risk Analyzer. Carnegie Mellon University, School of Computer Science, Institute for Software Research International, Technical Report CMU-ISRI-04-101.

Carley K.M. (2003). "Dynamic Network Analysis," in *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, Ronald Breiger, Kathleen Carley, and Philippa Pattison, (Eds.) Committee on Human Factors, National Research Council, National Research Council. 133-145.

Freeman L.C., (1979) "Centrality in Social Networks: Conceptual Clarification." Social Networks 1: 215-39.

Graham, J., Gonzalez, C., & Doyle, M. (2003) Using Communication Patterns in the Design of an Adaptive Organizational Structure for Command and Control In Proceedings of the 2003 Human Factors & Ergonomics Society, Denver, CO.

Graham, J., Schneider, M., Bauer, A., Bessiere, K., & Gonzalez, C.(2004) Shared Mental Models in Command & Control Organizations: Effect of Social Network Distance. In Proceedings of the 2004 Human Factors & Ergonomics Society, New Orleans, LA.

Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Human mental workload**,** 139-183. Amsterdam, The Netherlands: Elsevier.

Krackhardt, David. 1987. “Cognitive Social Structures.” Social Networks 9: 109-134.

Liang, D. W., Moreland, R. L., & Argote, L. (1995). Group versus individual training and group performance: The mediating role of transactive memory. Personality and Social Psychology Bulletin, 21, 384-393

Moreland, R. (1999). Transactive memory: learning who knows what in work groups and organizations. In L. Thompson, J. Levine, & D. Messick (Eds.), Shared cognition on organizations (pp. 3-31): Lawrence Erlbaum Associates

Nohria N, Eccles RG. (1992) Networks and Organizations: Structure, Form and Action. Boston, Mass.: Harvard Business School Press.

Obradovich, J., Graham, J., Schneider, M., Gonzalez, C. & Harder, R.(2004) Experimentation for Envisioned Worlds: Limitations of the Military After Action Review as Experimental Data, In Proceedings of the 2004 Human Factors & Ergonomics Society, New Orleans, LA.

Podolny, J & Page, K. (1998) “Network Form of Organization,” Annual Review of Sociology 24, p. 57;

Scott, J., (1992) Social Network Analysis. Newbury Park CA: Sage.

Strang, David and Nancy Brandon Tuma. 1993. “Spatial and Temporal Heterogeneity in Diffusion.” American Journal of Sociology 99: 614-639.

Wasserman, Stanley and Katherine Faust. 1994. Social Network Analysis: Methods and Applications. Cambridge University Press

Zsambok, C. E., (1997) Naturalistic Decision-Making Research and Improving Team Decision Making. In Naturalistic Decision Making. Zsambok, C.E. and Klein, G. Eds Lawrence Erlbaum Associates, New Jersey.

**Dynamic Network Analysis of the Network-Centric Organization:
Towards an Understanding of Cognition & Performance**

Chapter 4

Military C2 Experiment Collection & Database

Based upon initial success in framing the data collection problem and a solution during a pilot study, the US Army Battle Labs allowed me to lead a collection team in three successive studies. The method and the data have also been submitted to the US Army as an alternative to the NATO Code of Best Practices approach. The three studies include: Experiment #1: Ft Leavenworth Kansas, October- 2003; Experiment #2: Ft Lee Virginia, February-2004; Ft Knox Kentucky, June-2004. All of the experiments were part of the Unit of Action experiment series.

3.1 The Unit of Action Experiment Series

All of the experiments were conducted by the Army for a single purpose. They were all commissioned to gain a better understanding of the Unit of Action (UA) Command & Control Structure. The goal is to make decisions about the manning, automation support, and organizational design of the C2 Structure. The results are used in operations research (ORSA) models and presented in congressional briefings to justify manning and budget requirements. The findings are also used to write doctrinal manuals that the Army will use to guide organizational behavior (figure 3-1).

UNCLASSIFIED

Unit of Action Force Design

FOR USE IN SIMULATION

1ST UA

INITIAL Draft
Master
Battle Command Architecture
& Simulation Battlebook

INCREMENT 1 THRESHOLD
RESOURCED

17 May 2004
Based on Force Design of 30 June 2003

UNCLASSIFIED

U.S. Army Training and Doctrine Command

Figure 3-1. Iteratively updated doctrine manuals are one outcome of the UA experiment series.

Throughout the experimental series, the organizational design under testing was constantly updated. Each experiment resulted in a re-drafted Unit of Action architecture manual. Further, each experiment identified tests and conditions that needed to be applied to the next experiment. As a result, both the experiment design and the Unit of Action design were under constant revisement.

The Unit of Action

The Army has, up to now, deployed forces in 2,500 to 4,200-soldier Brigade Combat Teams. These consist of a ground-maneuver brigade (most divisions have three) augmented by other units, such as artillery battalions, which are controlled by the higher commander. The UA is a new “brigade based” structure that will replace the current arrangement, designed for the Cold War when the Army was prepared to fight giant set-piece battles on European soil, and the support roles were organized at the division level.

Brigade combat teams will be restructured into Units of Action. Although the exact configuration of units will vary, the Army has identified a basic Infantry UA design. The Infantry UAs will consist of approximately 3,000 soldiers that include combat, combat support and combat service support functions.

Beyond just new equipment, the UA has far better command and control. Command posts are standardized and integrate enabling capabilities and specialties into command post groupings. Headquarters manning is more robust, experienced, and knowledgeable than in current brigade organizations. Manning is robust enough for 24/7 sustained operations. The staff is more experienced, and enhanced with expertise it did not have -- especially Aviation, PSYOPS, Public Affairs, and Civil Affairs. Attached liaison parties from the Air Force, and from other services and SOF will be more robust. Enhanced battle command networks and functions speed informed decisions, coordination, and execution. Embedded and protected communications nodes insure robust and reliable communications.

While these experiments are continuing, the Army is not waiting until 2010 to begin the transition. The Army actually began this change in 2004 with movement to brigade combat teams where the brigade commander will have everything he needs to execute the operation. The 3rd Infantry Division, based at Fort Stewart, Ga. and moving into Baghdad as I write this, moved to four brigades as the Army's modularity test bed shortly after it returned from Iraq in 2003. The Army is standing up an additional two division brigades within the year and grow from 33 active-duty brigade combat teams to 48 by 2007. Therefore, the support Carnegie Mellon University provided these experiments is being put into practice, in battle, now.

The Unit of Action C2 Structure

The experiments reported in this dissertation were not conducted on the entire three thousand soldier UA. Instead, the focus is on the senior leaders of the proposed Unit of Action Command and Control element. That is an element that only makes up approximately fifty of the three-thousand soldiers in the UA. These fifty positions are absolutely critical as they are the few that maintain a top-level view of the organization and interface with the units above and adjacent to the organization. Their primary role is the UA commander's staff for all operations, intelligence, combat support, personnel, logistics and all combat service support. They are the integrators, coordinators, and planners for everything that happens within the UA.

Each one of the Battle Laboratory experiments had different organizational staff structures under testing. However, they were all fairly similar in concept. Each had a Mobile Command Group(s), Command Integration Cell(s), Maneuver Cell, Logistics Cell, Information Surveillance Reconnaissance Cell, and Effects Cell (table 3.1). While positions and total number of personnel changed in each experiment, these Cell concepts were stable throughout.

UA C2 Element	Description
Mobile Command Group (MCG)	Responsible for the overall mission. The MCG contains the unit commander, deputy commander and a small group of staff to support the commander.
Command Integration Cell	Responsible to integrate and double-check all activities of the

(CIC)	staff cells. The CIC contains a representative of each of the different functionalities of a Brigade (maneuver, fire support, intelligence, logistics, etc)
Maneuver Cell	Responsible to coordinate the activities of combat forces and combat support forces across the Unit of Action. This cell is oriented on operations.
Information, Surveillance, & Reconnaissance Cell (ISR)	Responsible for intelligence and communication related functions in the UA. Oriented on both facilitating the UA communication and inhibiting the enemy communication.
Sustainment Cell	Responsible for all logistics and combat service support in the UA. Includes supply, maintenance, personnel, and medical.

Table 3.1 Basic Unit of Action cell composition.

A Unit of Action Experiment

To understand and improve the Unit of Action design, the Army is not conducting experiments in the traditional sense. The experimental objective is to create an environment that reflects as closely as possible the conditions expected in battle. There are no experimental & control groups, independent and dependant variables, or treatments.

Instead, the experimenter brings in seasoned military professionals to role-play the required positions. They are given a set of computer based collaboration tools to facilitate coordination (Figure 3.2). A dynamic computer-based simulation is set up to represent blue, red and gray forces on the battlefield. A second set of role-players are brought in to simulate and control the enemy within the simulation based on a broad set of operating parameters and through free-play. Finally, the simulation is started and, periodically, stopped to collect observations from the seasoned military serving as role-players.

Figure 3.2. A typical cell in an Army Battle Laboratory Experiment configured with supporting collaboration equipment.

As a result, most of the findings in the experiments tend to be subjective. Also, much of the initial phases of the experiment are spent in team building and collaborative tool mastery. Lastly, it is not uncommon to change the simulation midstream to get at some concept the chief experimenter is interested in. In short, the experimenter is often suffering from as much uncertainty as the participants!

The value in these experiments is their validity to the current leaders of the force. The Army believes in the findings presented by seasoned professionals, not PhDs in white lab coats. After the experiment, through the discourse process, the seasoned military professionals work out what lessons they learned, which lessons were artifacts of the experiment, and which lessons are relevant to the future UA design. These lessons are then written up and translated into the Army's next doctrine.

Modified Unit of Action Experiment

After the pilot study (see chapter1), the seasoned professionals were intrigued about my findings that there was evidence of "hidden" organizational behavior their methods were missing. In particular they were concerned that they were not seeing all of the organizational behavior. The network behaviors and free-flowing collaboration were preventing a bird's eye view of the organization. As a result, I was invited to lead a team to collect and report at the following three experiments.

I sought to influence the experimental procedures by increasing the instrumentation. This required extensive participation in the experiment preparation, technical and managerial aspects. Much of the preparation is actually political in nature. Politically, there are many objectives and many stakeholders in each experiment. Each experimental objective has a stakeholder with their own data collection requirements. Two weeks prior to the experiment, the stakeholders meet to negotiate compromises or determine which collection is to be dropped.

Technically, the collection is software driven. An automated questionnaire is absolutely necessary to provide, within an hour of collection, feedback and data analysis to the chief of experimentation. These collection and analysis tools have to reside on the same platform and network supporting the simulation. As the network is typically not finalized until the week before, the final set of collection software cannot be loaded and debugged until the week before the experiment.

Managerially, a collection team has to be on the floor at all times during the experiment. The participants are often overwhelmed by the sheer amount of information being collected on them during the experiment. The collection instruments will not be completed without a human team to support questions, definition explanation, and even chiding of the participants. Further, during the experiment, there are often changes of direction in the experimental goals that the team has to read and use to alter the collection schema.

During each experiment, I had a team of four to six aids working on the collection team. In some instances, the team was all Carnegie Mellon University students and in other instances I was given partial personnel support by the experiment director. In every experiment except Ft Knox, the software technician and network installation was done by Mr Mike Schneider. I personally led all collection objective negotiations with the experiment director and stakeholders.

In exchange for participation, my team was required to provide a broader level of analysis than my own research objectives to support the Army's Unit of Action design process. As such the experiments contain a lot of peripheral activities and collection instruments not directly applicable to this dissertation. Further, each experiment built off of the previous work. This iterative process results in very rapid prototyping, but it is not the optimal conditions for formal hypothesis testing. However, the result is that, in addition to having new theoretical findings in Shared Situation Understanding, the findings are actually in a usable form for existing organizations.

3.2 Experiment #1 Fort Leavenworth Kansas, October-2003.

On 3 October 2003, 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 3.3). 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.

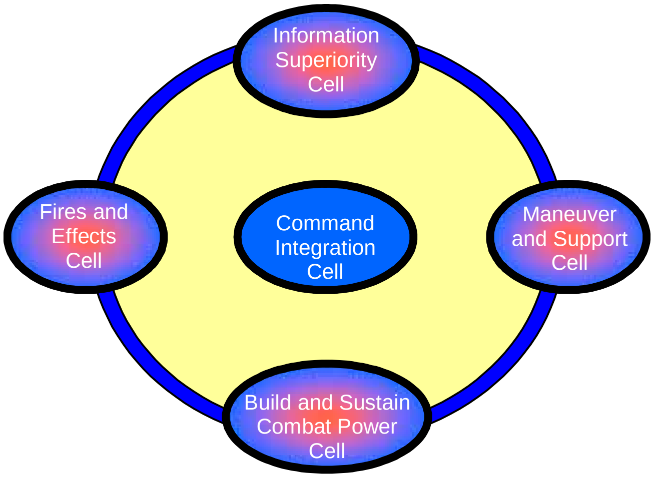


Figure 3.3: Organizational Structure for the BCBL-L exercise. Each Cell consisted of 3-8 agents that could flexibly organize and collaborate as necessary (see table 3.1 for Cell explanations).

Each cell member was not limited by their cell membership and was free to coordinate with other cell members as necessary. The four nodal cells, shown on the outer ring of figure 3, consist of functional groupings of staff roles. The Fire and Effects Cell is responsible for indirect fires and psychological operations. The Maneuver and Support Cell is responsible for ground maneuver and protective activities. The Information Superiority Cell is responsible for the intelligence collection and communication support. The Build and Sustain Combat Power Cell is responsible for logistics. Each of the cell members were expected to identify problems and conduct intra and inter cell collaboration as the battle scenario unfolded.

Data Collection

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.

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.

I 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. Operational workload congruence 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 3.4). 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.

Workload Questionnaire

Low 1 2 3 4 5 6 7 High

Please answer the following questions about WORKLOAD.
Rate Each Question From 1 to 7 (1- lowest, and 7 - highest).

RATE YOURSELF

1 MENTAL DEMAND
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

2 TEMPORAL DEMAND
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

3 EFFORT
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

4 OWN PERFORMANCE
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

5 FRUSTRATION LEVEL
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

6 PHYSICAL DEMAND
○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

Page 1 of 14

Figure 3.4: Workload Questionnaire Administered During the Fort Leavenworth Exercise

Mental model congruence was calculated 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.

A copy of the technical report submitted to the Fort Leavenworth Battle Laboratory is contained in Appendix 4. The technical report contains organizational design recommendations based on an ORA analysis and comparison of the proposed organization and the actual organizational behavior. Experimental results relevant to the study of situation awareness are contained in chapter 4.

3.3 Experiment #2 Fort Lee Virginia, February-2004.

In February 2004, the Ft Lee 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. They were purely interested in the logistics activities associated with the Unit of Action staff. As such the majority of the participants had similar logistics background and training. Further all of the tasks injected into the experimental scenario were designed to stress logistics play in the simulations.

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. During their scenario, artificial barriers were in place, 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 (chat, email, etc). A plan-execute-plan-execute cycle was used in the simulation.

All people could see all other people.

There were a total of eight cells in this logistics organization (Figure 3.5). The UA-MCG (Mobile Command Group) consisted of the commander and immediate staff who are responsible for all critical decisions on unit maneuver. Six of the nodes were divided into two teams of three cells. One team focused on aviation logistics (CAB-MCG, CAB-BSC, & AVN-BSC) while the other focused on maneuver logistics (UA-CIC, UA-MSC, & UA-BSC) . The FSB, forward support battalion owned all of the logistics resources necessary for executing logistics support.

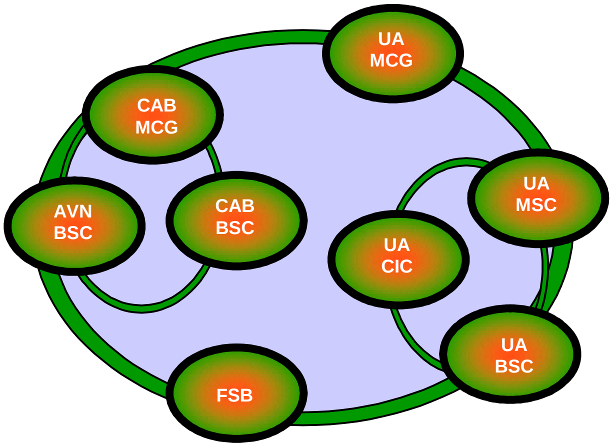


Figure 3.5: Logistical Unit of Action Organization Design.

Data Collection

Observations and data collection were conducted over two days immediately following a one-week training period. A plan-execute-plan-execute cycle was used. Data was collected every 60-90 minutes using networked questionnaire that asked the participants for feedback regarding their prior session. Questions included communication frequency, top three tasks worked on, NASA-TLX on self, and NASA-TLX estimation of others. Questionnaire data was collected for a total of 6 sessions of approximately 70 minutes.

Data was collected every 60-90 minutes using networked questionnaire that asked the participants for feedback regarding their prior session. Questions included communication frequency, top three tasks worked on, NASA-TLX on self, and NASA-TLX estimation of others. Questionnaire data was collected for a total of 6 sessions.

A copy of the technical report submitted to the Fort Lee CSS Battle Laboratory is contained in Appendix 5. The technical report contains organizational design recommendations based on an ORA analysis and comparison of the proposed organization and the actual organizational behavior. Experimental results relevant to the study of situation awareness are contained in chapter 4.

3.3 Experiment #3 Fort Knox Kentucky, June-2004.

The Unit of Action Maneuver Battle Laboratory (UAMBL) gathered two hundred- twelve experienced army officers to serve as participants in an integrated Unit of Action experiment. Positions ranged from platoon level to division level command and staff groups. The participants were distributed at six locations throughout the United States (figure 3.6) and connected via voice communications, chat, email and shared whiteboard. The cell members operated from mockups of mobile command and control vehicles.

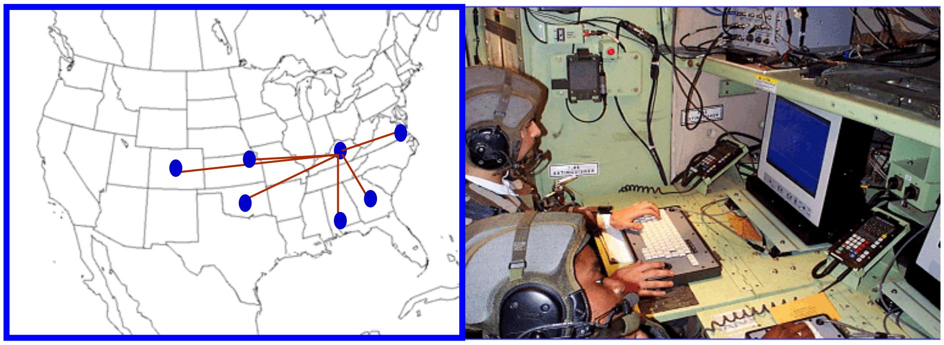


Figure 3.6. The Ft Knox experiment conducted with participants distributed across the United States and positioned in mockups of mobile command and control equipment.

The participants gathered information, coordinated with appropriate staff members, and entered battlefield actions into the simulation. A plan-execute-plan-execute cycle was used with the focus on the UA command and control staff. This was the most robust test of the UA design as there were role-players in multiple subordinate and higher echelon units.

This was the first experiment to test the two command integration cell (CIC) concept (figure 3.7). The two cells were originally designed to provide a 24 hour battle coordination capability. However, during the experiment, the two CIC concept morphed into a planning and execution focus with each cell taking a different responsibility.

.

Figure 3.7. Unit of Action Staff in two CIC configuration.

Data Collection

Observations and data collection were conducted every four hours over fifteen days immediately following a eleven day training period. 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

Questions included communication frequency, top three tasks worked on, top threats to the operation, NASA-TLX on self, and NASA-TLX estimation of their commander. Questionnaire data was collected for a total of 18 sessions.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, I constructed a social network graph for every session. The social network graphs were used to calculate the social network distance (geodesic) between each player.

During this exercise participants only completed workload rating for themselves and their commander using the NASA-TLX. Due to limitations on the collection software, I could not randomize a rating process as in the previous experiments.

I therefore elected to create an SA congruency measure. This 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, I constructed a pair-wise measure for each pair of participant responses, which estimated the similarity of their risks. This provided a measure of risk congruency of similarity.

A copy of the technical report provided to the Ft Knox Unit of Action Battle Laboratory is not included in this dissertation as some elements are considered sensitive. However, it was constructed using an ORA based analysis of the agent, task, knowledge meta-matrix of the experimental organization's behavior. The results related to this dissertation are contained in chapter 4.

3.4 Comparison of Data Sets Collected Across All Experiments

Figure 3.8 is a visualization of how the data collected during each experiment maps onto Endsley & Jones' (1999) shared situation awareness theory. Note that all three experiments are necessary to understand all of the components of Shared Situation Awareness theory. Shared SA Mechanisms had to be revisited in every experiment to get a firm understanding of the processes. However, success in understanding other elements of shared situation awareness theory in early experiments, allowed the future experiment objectives to shift to other areas.

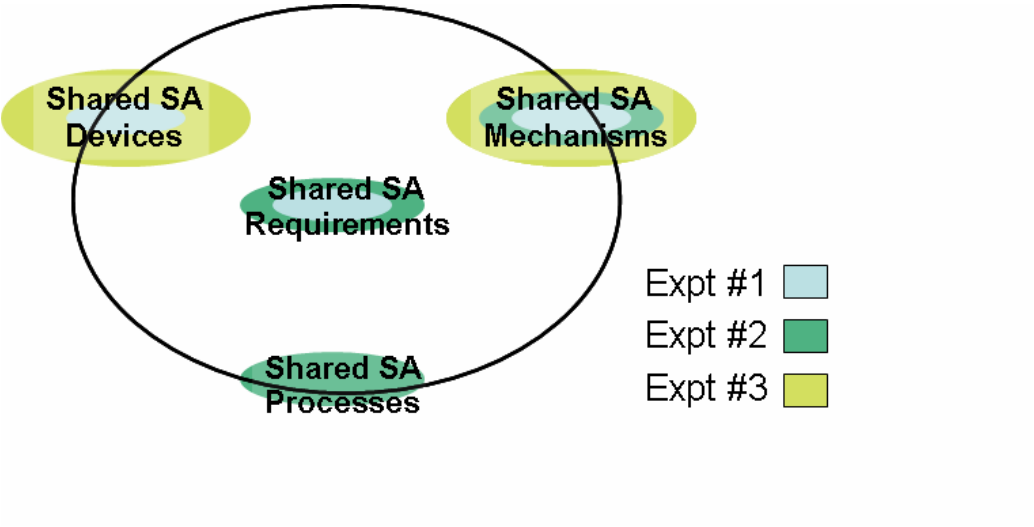


Figure 3.8. Experimental collections as related to shared situation awareness theory.

Table 3.2 is a matrix of each experiment against each of the questionnaires and measures administered. The experimental design or negotiations with other experimenters often limited the instruments and measures. However, I used a wide enough range of measures across multiple experiments to obtain the desired understanding of shared situation awareness.

Table 3.2. The three experiments matrixed against the collection and measure technique applied.

Discussion

I entered into these experiments to collect data for my dissertation. However, it is becoming evident that Carnegie Mellon University's participation in these experiments is changing how experiments are designed and even how the Army thinks about organizations. Experiment design now consistently includes measures of the network and hidden leaders. The Army's top thinkers at the Center for Army Analysis at Ft Belvoir, VA has requested briefings and demonstrations for their analysts. Last, the Army's chief experimenters now expect our participation in current and future organizational design analysis. In short, bringing a dynamic network perspective to these experiments has had a great immediate impact on the entire Army.

There are two important deficiencies in these experiments: 1) the duration of the time period between collections, 2) the field measures of shared situation awareness.

Primacy Recency.

Insert Discussion from Graham et al HFES 04 here

Field Measure of Situation Awareness.

Write up the noisy measure issue here

Given my thesis goals this research depends upon these unique databases. First, these are clearly network organizations. Most of the network organization literature focuses upon organizations that are formally hierarchical with hidden or informal network-type behavior. These experiments were conducted on organizations formed to specifically behave as a network.

Second, the members of the prototype organizations were all selected for their functional expertise. The most common complaint about Situation Awareness research is the use of novices that never obtain proficient skill sets to achieve SA during the laboratory experiment. However, in these experiments, the participants are not untrained college sophomores in a ten dollar an hour experiment, but instead the participants have up to thirty-five years of functional expertise and experience of operating as a command and control team member. Each member knows their roles, as a result, the only new skill required is operating within a network, instead of strong hierarchical, organization.

Conclusion

Clearly, these databases are unique and will support an iterative examination of shared situation awareness as well as the factors that mediate network change. These databases are unique in that they contain the essential data to construct a network at multiple times in the organizations life. Further the experiments use extreme forms of network organizations assembled specifically to test the network organization concept. As some measures are the same across all experiments, I am able to test metrics developed in the previous experiment. However, the iterative experimentation also allowed for the development of new field measures of shared situation awareness components that were identified during the previous experiments.

Dynamic Network Analysis of the Network-Centric Organization: Towards an Understanding of Cognition & Performance

Chapter 3

The Military Command & Control Experiment

This research is the first to use field data sets of temporal organization communications, tasks and training to measure, understand, and describe shared situation awareness. To this point, there has been no experimental (field or laboratory) work, to date, on the relationship between shared situation awareness and communication network evolution. This is due, in part, to the difficulty of large organization network data collection, limited understanding of dynamic network studies, and limited field measures of shared situation awareness. Further, there are simply no suitable databases available to study and describe time-based network evolution and shared situation awareness evolution. As a result, this thesis proposal includes a series of field experiments to develop an appropriate dataset. The data sets will be used to develop a shared situation awareness metric (chapter 4) and model network evolution (chapter 5).

The US Army provided an opportunity to integrate into their experiments to collect data. The military is undergoing a major transformation in the way it equips, organizes and trains. The Army battle laboratories have been charged with conducting the experiments to test and validate the technologies, organizations, and roles that are will define the future force. The fifteen battle laboratories, located throughout the United States, spend between one and fifteen million dollars per experiment. Two to eight experiments are conducted per year.

The experiments fit my dissertation requirements as the context and collection provides:

- Nodal organization design

- 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)

These experiments are neither laboratory experiments in the academic psychology tradition nor are they field experiments in the tradition of management science. First a select group of experts design the prototype organization. Once a prototype organization has been identified for testing, a battle laboratory is selected to conduct experimentation. The selected battle laboratory director is given a set of tasks the organization is designed to accomplish, the organization design, and a list of roles in the organization. The battle laboratory director then designs a three to fifteen day experiment to test the organization.

First, the experiment is set up to replicate the placement of the prototype organization. The elements that will control the prototype organization and the elements that will be subordinate to the prototype organization are included in the design. As military uses a single chain of command, there is only one controlling element. The controlling element is replicated by a leader with a few staff positions to give direction and stress the prototype organization. The chain of command system allows for multiple subordinate elements. These are replicated by commanders and staff role-players two levels down from the prototype organization under testing.

Each role-player in the prototype organization and external elements are filled by a participant charged with preparing, role-playing, and reflecting on a position in a hypothetical organization of the future. The role-players/participants are selected for their experience in the military and how close their specialty in the current force is congruent with the roles of the future force. The goal is to have each position filled by a soldier that has more experience than is called for in the role. This allows the soldier to reflect on their own performance as it relates to previous experience.

The workspace for each role-player position is created using the best available replications of future technologies. The minimum requirement is for a visualization of the simulation and a method to electronically communicate and collaborate with the other role-players. In some cases, state of the art chat, email, whiteboard, and voice over internet protocol is provided to support collaboration. In some experiments all communication is face-to-face.

A computer-based simulation serves as the stimulator for role and organizational activities in a military C2 context. The simulation is set with initial start conditions and behaviors to replicate future capabilities. The large group of role-players (28-221) then takes positions in their workstations and behaves, coordinates, and makes decisions with the other role-players in their organization.

The experiment data consists of qualitative observations from the role-players and quantitative collections from the simulation. Since the participants are experienced military officers, their reflections and answers to open-ended surveys are considered knowledgeable insights to how well the organization structure and role placement in the organization structure supported the mission. These observations are collected and then collated into findings for the organizational design. The amount of time an action takes is collected as quantitative data from the computer simulation. The premise is that if a decision takes a shorter amount of time, it is a better decision.

The data analysis method is based on a “bucket” system. The lead experimenter will designate a set of general areas in which he is interested in learning about the organizational design and label them as “buckets”. The participants and observers are then responsible for adding their comments and observations to the bucket that is most relevant. At the end of the experiment, a senior group of analysts determine how to assemble the observations into brief-able sets of findings about each “bucket” area.

The results section is then organized by issue. The issues are predetermined based on a set of research questions the leaders have about the new organizational form. Each issue has an information requirement, an emerging insight, supporting evidence, and a recommendation. The information requirement comes from an approved document of new capabilities the military want the organization to exhibit and provides a numbering reference system for each finding. Emerging insights are a collection of analysis that indicate a particular hypothesis may be the case. Supporting evidence is a list of data points used in the analysis and emerging insights. Lastly, recommendations consist of changes to the organizational design or changes to future experimental designs.

The stimulation for role-player behaviors is the interaction with other members and the interaction with the simulation. The methods to interact with others varies across experiments. Some experiments allow face to face discussions, while other experiments force coordination to occur through collaborative technologies. The collaborative technologies include telephone, radio, voice over internet protocol (VOIP), chat, email, and whiteboards.

The simulation includes representations of all potential actors in a battlefield environment. The computer based visualization enforces time, distance, and space relationships between all actors in the environment. Each workstation visualization only provides data on the actors that would be available from the activities of the actors. For instance, a maneuver commander can only directly know about enemy that he or his units have made contact with.

Codified Military C2 Experiment Process

In 1999, the North Atlantic Treaty Organization issued a technical report laying out a “Code of Best Practice” for command and control assessment. The concept was to provide a framework for conducting Command and Control (C2) experimentation. The sense was that too many resources were being devoted to experiments that were providing little return on investment. The Code of Best Practices, therefore, offered the following guidance to conducting C2 experimentation.

1. The experiment must be based on a clear set of issues of interest. Issues of interest are developed from a general articulation of the problem to be addressed. (Good Games Paper).
2. The experimenters must identify, in advance, the organizational and cultural issues of the role-players (values, behavior, decision processes) that will impact the experiment results.
3. Scenarios must be designed that prompt the role-players to perform activities related to the issues of interest. The scenarios become the drivers of behavior.
4. In advance of the experiment, Measures of Merit must be established. Measures of Merit determine what data should be collected during the experiment. The assessment of that data is the Measure of Merit as related to a issue of interest.

While the Code of Best Practices has had a revolutionary effect on the military experiment community, its implementation has been suboptimal. The four parameters of best practice have proven to be too broadly defined and too removed from the academic scientific process. Without concrete procedures, the experiment implementations have been weak. In fact, in 2004, NATO commissioned a committee to review how the Code of Best Practices was implemented and how it should be improved.

Problems with the Military C2 Experiment

These experiments do not have hypothesis. While there may be some general expectations for how the organization will perform, those expectations are not operationalized into hypotheses. As a result there are numerous issues related to validity of the experimental results. Without a hypothesis, there are no conditions to accept or reject the findings. There are no linkages made, in advance, between the data collection and the expected outcomes. Instead, data and observations are collected and summed into findings for hypothesis generated after the experiment.

There is no baseline of performance and behavior in the experiment. The typical experimental data is not compared against similar data collected on a control organization design. Instead, subject matter experts evaluate the data and observations for departures from normal behavior as defined by their experience. Any departure from the SME’s intuitive expectations are considered significant data points and findings.

As a result, unless the same SMEs attend every experiment, only a cursory comparison of results across different experiments is possible. As, there are no standardized collection instruments across the experiments, every experiment develops a new set of collection instruments to collect data. Analysts at each experiment develop new collection instruments based on the objectives of the experiment, the set up of the organization, the number of observers available and other experimental design factors.

Few of the instruments are pilot tested. A pilot test involves distributing or conducting the collection instrument on a representative sample of people ahead of the experiment. The pilot test will determine the mechanics of collection tool deployment, the range of typical data the collection tool will obtain, and how the data is best analyzed. Without a pilot test, the collection tools are often changed multiple times through out the experiment as problems with deployment, data, and analysis are determined.

The collection instruments have a high dependence on the memory of the participants which leads to missed and skewed data. Most collection tools and interviews are conducted after a six to eight hour block of the scenario. Obradovich et al (2004) found that military experiment role-players, questioned at the end of the day, forgot up to sixty percent of the key activities of each day. In fact, the role-players, on end of day collection instruments demonstrated a strong primacy-recency effect by best recalling events that occur at the beginning and at the end of the day. The participants were more likely to forget events that occur in the middle of the day.

Little Motivation for Change

Despite the problems with the scientific method, this form of experiment is very popular with the military. Ultimately, the Army believes in its people. The participants are experienced military officers that have useful insights to how the Army works. They can provide meta-analysis of their own behaviors to contribute to the experimental outcomes. Philosophically, the best equipment or design if not accepted and embraced by the experienced role-player is not useful.

The scenarios are set up to reflect the most realistic conditions possible. These scenarios are not like university laboratory experiments that use contrived games and stimulations to elicit behaviors from the participants. In fact, experiments in over-controlled environments are distrusted by the military. Instead the scenarios reflect the complexity and dynamism indicative of a military operation. The organizational design, therefore, is stressed in conditions that reflect the best estimation of future military operations. The resulting organizational performance is considered a reliable replication of what our soldiers will have to contend with.

The experiment objectives can be adjusted during the experiment. If it is determined that an objective was already met or cannot be met, the experiment director can reassess the experiment objectives. Given that the experiment cost can range from two to fifteen million dollars, the lead experimenter has the leeway to make adjustments that will return the required findings and get maximum return on investment.

The experimental sponsoring agency can control how the findings are published. These experiments are often used to justify major funding decisions on weapons, computer and training systems for the Army. If left unchecked, stakeholders can sway the analysis to support the funding decision they are interested in. Under this structure, the experimental sponsoring agency can control what findings are legitimate, robust and releasable.

Recommended Changes to the Military Experiment Design:

Every military experiment needs a set of hypotheses. Using the experimental objectives, the lead experimenter should develop research questions and hypothesis of expected behavior. The hypothesis should be based on how a current organization behaves, the differences between a current organization and the prototype organization, and how the differences are expected to change role-player and organizational behavior.

All data gathering should be based on these hypotheses. The collection instruments should be targeted to reject or accept the hypotheses. Armed with the right data, the analyst can make an objective case for why an organizational design did or did not perform as expected. Subject Matter Experts can then make a subjective, experience based case for why the organization behaved as indicated by the data.

Data should, as much as possible, depend on the behaviors of the roleplayers. The role-players' own meta-analysis of their behaviors is valuable, but can be skewed by recall. The role-players should be valued and studied as experts performing in a new context. Their new behaviors in the new organizational context provides an objective delta for analytical comparison and is not skewed by recall issues.

Collection instruments should be standardized and take advantage of the power of the computer. Standardized collection instruments allow for comparison of results across experiments. This ability leads to the ability to compare different prototype organizational designs undergoing incremental change. A standard set of data allows comparison of behaviors across different type of organizations (UA, UE, etc).

Lastly, the collection instruments need to take advantage of aggregation ability of computers. Each experiment has different levels of analysis (individual, cell, team, complete) of the prototype organization. The inclination is to develop collection tools that target that level of analysis. However, by collecting at the individual level, computers can aggregate the results to the desired level of analysis (cell, team, etc). For instance, with individual workload data, the computer can develop cell workload, team workload, and overall organizational data. The result is a single collection tool for all experiments but tailored analysis for the each experimental objectives.

Dynamic Network Analysis of the Network-Centric Organization:
Towards an Understanding of Cognition & Performance

Chapter 7

A Network-based Shared Situation Awareness Metric

While the situation awareness literature base is rich and complementary, shared situation awareness has only been the specific research topic of two studies (Endsley & Jones, 1997, 2001). Therefore, there is, as of this time, no 'right' or accepted method to conducting research and measurement of shared situation awareness. As a result this thesis proposal will, necessarily, develop a novel approach to studying shared situation awareness.

Using the three field experiment databases (chapter 3), I will develop and implement a Shared Situation Awareness performance metric in Organizational Risk Analyzer (ORA). First, I will decompose shared situation awareness theory and then iteratively improve the shared situation metric using the three prototype organization experiments. While figure 5 explains the measurements contained in each of the field experiments, figure 4.1 is a visualization of the decomposition approach of Shared Situation Awareness.

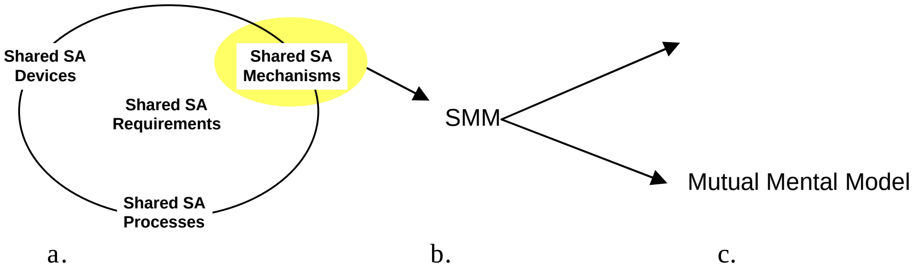


Figure 4.1. Decomposing shared SA into its literature based components. a. shared SA, b. shared mental models, c. components of shared mental models

4.1 Iterative Metric Development

Each decomposition of shared SA begins with a literature based description of each factor. For instance, Shared SA can be broken down further into its four components (figure 4.1 a). Those four components are Shared SA requirements, Shared SA devices, Shared SA mechanisms & Shared SA Processes. For example, in Figure 4.1 above, a. shared SA mechanisms

(Endsley & Jones, 1997) description is strongly influenced by (b.) the literature on shared mental models (Salas et al, 1992). (c.) Shared mental models can be further decomposed into the common situation model and the mutual mental model (Entin & Serfaty, 1999). Selecting one of the two, mutual mental model, I can now develop dynamic network based measures that accurately estimate the mutual mental model.

The metrics will be integrated through an iterative process (figure 4.2). I will continuously develop literature supported network-based predictors of a component of shared situation awareness. The network-based predictor will be tested by implementation in Organization Risk Analyzer (ORA). The new ORA measure will then be initially validated against the collected field experiment data.

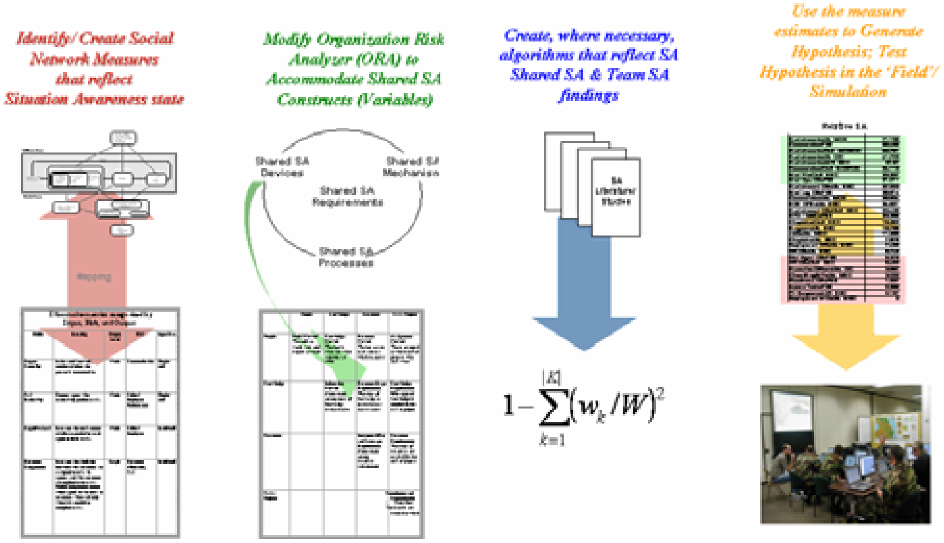


Figure 4.2. Iterative approach to Shared SA measurement development.

Metric validation is accomplished with the data set collected at each of the battle laboratories. For each experiment, I formulate a measure that reflects one of the four shared SA constructs and collect it along with the dynamic communication network. The dynamic network based estimate of shared SA state can then be validated against the measured or self-reported shared SA state.

4.2 Initial Metric Success

One example of this decomposition process has been submitted to NAACSOS '04 (Graham, 2004). The data was collected during Experiment #1 at Ft Leavenworth. The following is a description of the process and the product.

One critical component of Shared Mental Models is a mutual model of roles and tasks (Figure 4.1) (Entin & Serfaty, 1999). Mutual mental models are a common 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). Mutual mental models are not unlike transactive memory which is defined as, in a group or organization, who knows what information (Wegner, 1987; Argote, 1999).

Mutual mental 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 questionnaires that identify a participant's mission to develop an understanding of their mental model. An index is then calculated to determine the closeness of the two team members' mental models (Graham & Schneider, working paper).

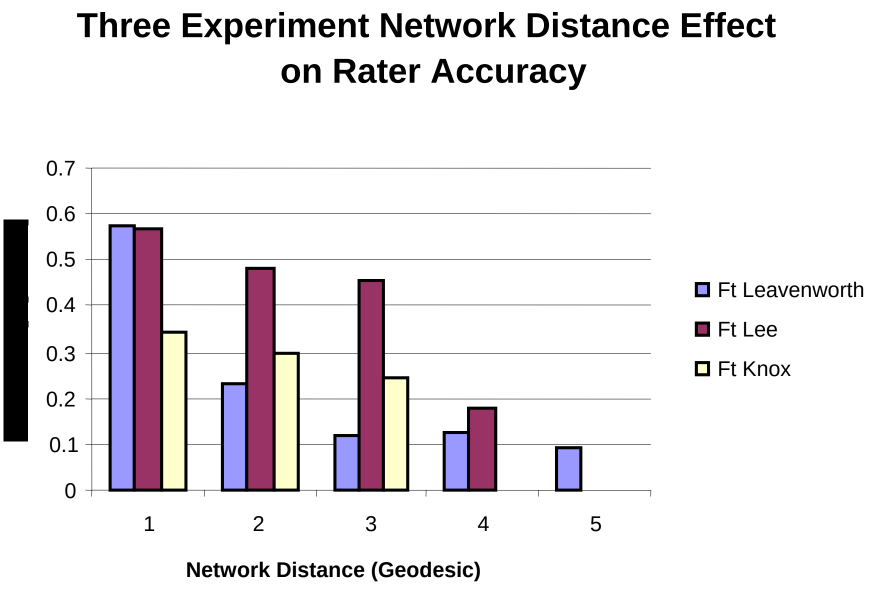
4.2.1 Network Dyad Approach

For the purposes of this initial shared situation awareness metric development, I only considered within each dyad, the physical proximity and the communication network distance as well as the overall distance from a high authoritativeness member.

4.2.1.2 Communication Network Distance

Communication, particularly in a command and control environment, also provides cross-training. 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, Schneider, Gonzalez, 2004). We define communication in terms of the number of nodes on the geodesic between two agents. In the case of Graham et al (2004), we found that the relationship between communication and mutual mental models was stronger than between physical proximity and mutual mental models in a C2 organization.

Figure x (below) is the network distance (geodesic) as related to rater accuracy in rating others workload for all three experiments. Note that in every case mean accuracy significantly decreases as the geodesic between any two dyads increases. (INSERT F-Test results here).



1. Trend for all three
2. Effect on Measure

3. Resulting component of the measure

$$SAm(mm) = \chi C_{ij}$$

4.2.1.1 Physical Proximity

Rouse, Cannon-Bowers, & Salas, (1992) found that cross-training can occur through two methods, physical proximity and communication. Findings by 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 shared mental models. In the case of Graham, Schneider, & Gonzalez, 2004), we found that physical collocation was twice as likely to produce a shared mental model (figure 4.3).

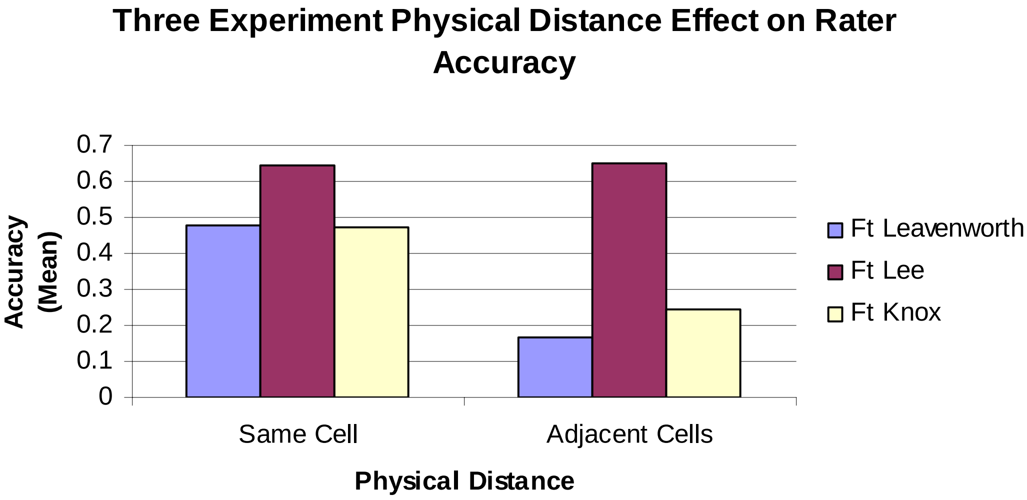
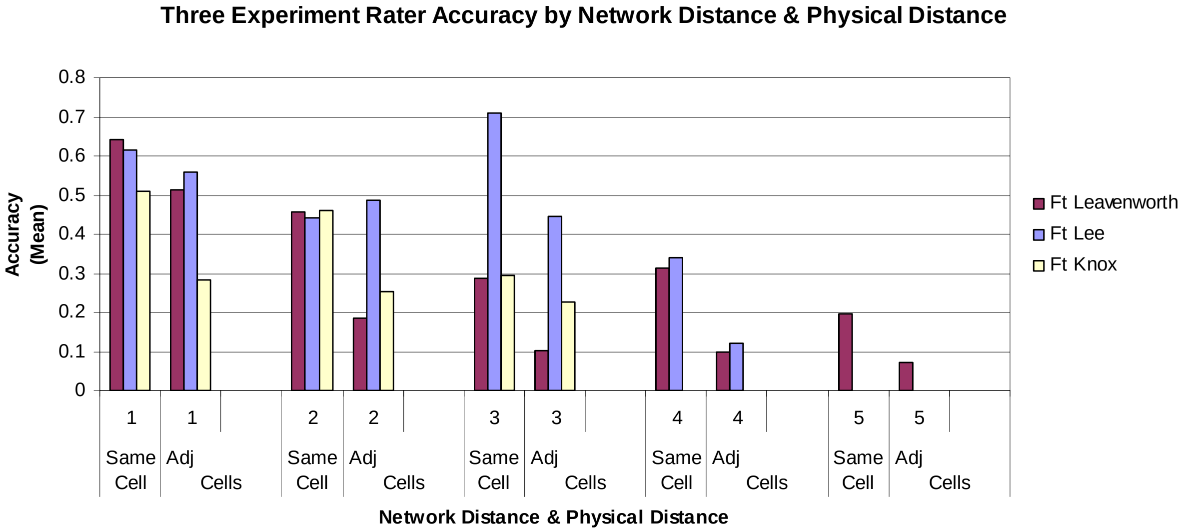


Figure 4.3: Mean shared mental model by physical distance (shared cell vs non-shared cell) (n = 3028) (Graham, Schneider, & Gonzalez, 2004)



4.2.1.3 Authoritativeness

In the command and control context, leaders tend to have the highest situation awareness (French & Hutchinson, 2002). Within a network, identification of informal and formal leaders can be achieved by measuring authoritativeness. An agent with high authoritativeness has nodes with high hubness values point to it (Kleinberg, 1998). Kleinberg's authoritativeness algorithm is a well-known link analysis algorithm that identifies "authoritative" or "influential" webpages in a hyperlinked environment. Intuitively, by thinking of a communication as a citation, a agent i is more of an authority (i.e. highly-referenced agent) as compared to agent j if there are more hyperlinks entering i from hub agents, where a hub is simply a agent that is a valuable source of communication links to other agents. Likewise, a agent i is a better hub than agent j if there are more communications exiting i into authoritative agents. Given a set of n agents, we first construct the corresponding n-by-n adjacency matrix A, such that the element in row i and column j of A is 1 if there exists a communication from agent i to agent j, 0 otherwise then iterates the following equations:

$$x(t+1) = AT\ y(t) = (ATA)x(t) \qquad (1)$$

$$y(t+1) = Ax(t+1) = (AAT)\ y\ (t) \qquad (2)$$

Where the i-th element of x denotes the authoritativeness of agent i and the i-th element of y denotes the value of agent i as a hub. With the vectors x and y initialized as vectors of ones and renormalized to unit length at every iteration, as t approaches infinity, x(t+1) and y(t+1) approach x* and y*, the principal eigenvectors of ATA and AA T, respectively.

Even though the Kleinberg algorithm is originally intended to locate hubs and authorities in a hyperlinked environment, hubs and authorities map very well to the agents and communications in a command and control environment. An agent with high authoritativeness will have a high degree of access to agents that are participating in critical organizational tasks. Further, their role as authorities starts them with high experience that is only strengthened over time. As a result agents high in authoritativeness are better able to maintain a mutual mental model. In fact, it is likely that agents with high authoritativeness can serve as the base-line mental model with respect to all other agents.

4.2.2 SA Mechanisms Integrating Metric

This measure of SA Mechanisms (SAm) seeks to integrate descriptors of agent to agent communication and collocation with an individual measure of authoritativeness.

For any two agents, SAm can be described as follows:

SAm = f (Authoritativeness , Physical Proximity , Social Distance)

$$SAm(mm) = \alpha A_i + \beta P_{ij} + \chi C_{ij}$$

I first define Authoritativeness based on Kleinberg (1998) description (see previous section). I also define proximity Pij in binary form. Given a set of n agents, we first construct the corresponding n-by-n adjacency matrix A, such that the element in row i and column j of A is 1 if agent i and agent j are collocated in the same cell, otherwise 0.

Lastly, I define communication in terms of the nodal length of the geodesic between two agents. Given the same set of n agents, we first construct the corresponding n-by-n adjacency matrix A, such that the element in row i and column j of A is 1 if there exists a communication from agent i to agent j, otherwise 0. We then calculate the number of links on the geodesic between agent i and j, g(i, j). To place communication into the same linear positive formula as authoritativeness and proximity, divide into 1.

4.2.3 Metric Validation

The SA Mechanisms algorithm was against a data set collected during an Army Command and Control Exercise. The data set is from a trial fifty-six army officer organization. Each officer serves as role-players for an experimental command and control staff that was put through a computer scenario for four days.

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 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. 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. 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.

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.

Shared mental models were 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. When rating other people, participants had the option of selecting “Don’t Know” for each of the seven questions. As the participants could opt out of an evaluation of others, ‘willingness-to-rate’ is used as a proxy for the state of the shared mental model. Willingness-to-rate was determined by number of the six NASA TLX items for which the participant chose to rate the other person divided by the total number of questions. A willingness to rate others index is an indication that the rater has confidence in their knowledge about the ratee’s capabilities, tasks, and organizational situation (Graham, Schneider, Bauer, Bettsiere & Gonzalez, 2004)

4.2.4 Results

The variables of interest are the SA Mechanisms algorithm and the measured shared mental models. For a point of comparison, we also calculated social network distance, physical distance (proximity), authoritativeness and betweenness centrality (Freeman, 1979) as alternative surrogate measures of SA Mechanisms (Table 4.1). In the regression model, session was used to control for the effect of learning in this new organization. Regression results indicate that the SA Mechanisms measure was the best predictor of Shared SA Mechanisms. Betweenness Centrality was not a predictor of Shared SA Mechanisms.

Parameter	**Estimate**	**Std Error**	**t-Ratio**
Intercept	-0.2842	0.08354	-3.4022*
Betweenness Centrality	0.0002	0.00008	1.846136
Authoritativeness	-1.808	0.46335	-3.90199*
Network Distance	0.0576	0.01853	3.108055*
Physical Proximity	-0.6505	0.08916	-7.29631*
SA Mechanisms Measure	0.9302	0.08034	11.57778*

Table 4.1. Network Measures of Shared Situation Awareness Mechanisms (*p<.01)

4.2.5 Discussion

SA Mechanisms, as modeled, had a moderately strong relationship with the measured shared mental models. The relationship became stronger through session eight at which time it leveled out r= ~ xxx . As the organization has just formed and the members met each other for the first time at the start of the sessions, we can attribute the rise in performance of our measure to a learning of the members’ capabilities, tasks, and situation. We

> can also infer that our measures are more reliable for a network that has stabilized than an organization undergoing significant learning.

The SA Mechanisms measure had an advantage of combining both agent-agnet communication and agent-agent physical proximity measures. Graham, Schneider & Gonzalez (2004) found that both social network distance and physical distance were predictors of mental model congruence. Further these distance measures were found to be qualitatively different from each other. Physical distance provides information about who is likely to use face-to-face coordination and who is proximate. Network distance is capable of providing information about who is linked by the task structure, context, and organizational structure. The weak correlation between the two distance measures further supports the difference between network distance and physical distance. 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.

$$SharedSA_{ij} = \alpha D_{ij}$$

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	$D_{ij} : DirectCommuni$
Model 2	$P_{ij} : PhysicalProximit$ $C_{ij} : Geodesic(SocialNetworkDist)$ $A_{ij} : NetworkLeader$
Model 3	$P_{ij} : PhysicalProximit$ $C_{ij} : Geodesic(SocialNetworkDist)$ $A_{ij} : NetworkLeader$ $H_{ij} : Homophil$

Figure 13: 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 and 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.

$$SituationModel_{ij} = \alpha A_{ij} + \beta P_{ij} + \chi C_{ij}$$

5.3.2.1 Model 2 Factors

<u>Physical Proximity</u>
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 (E. 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 14 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.

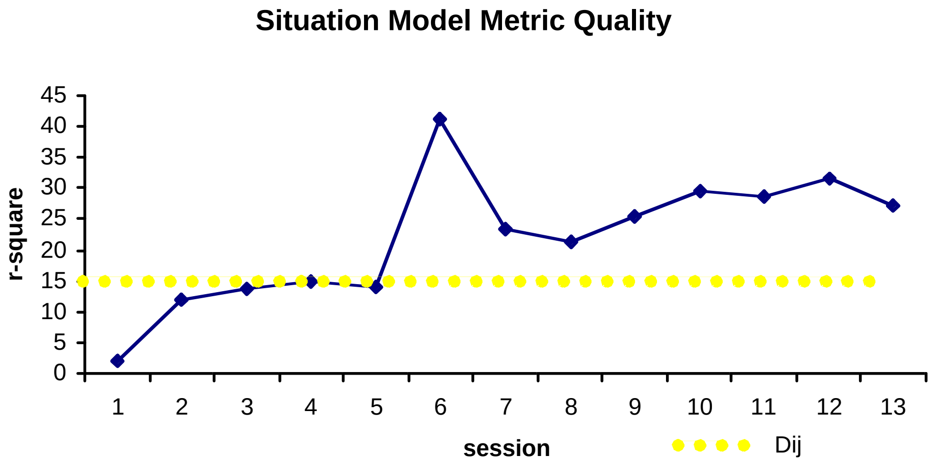


Figure 14: Comparison of the Baseline Model Against Model 2

Figure 14 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

$$SharedSA_{ij} = \alpha A_{ij} + \beta P_{ij} + \left(\frac{\delta H_{ij}}{\chi C_{ij}} \right)$$

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.

5.3.3.1 Validation of Model 3

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.

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 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.

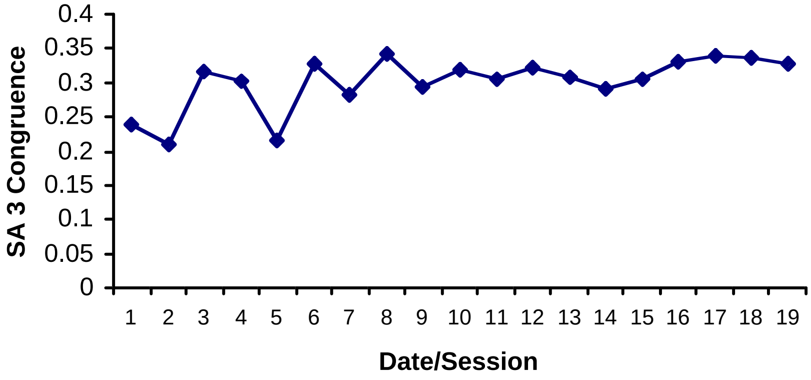


Figure 15. Model 3 - level 3 Congruence Over the Duration of the Experiment

Figure 15 is the mean situation awareness (level 3) congruence for the organization. In the early stages of the experiment, there were tremendous fluxuations in congruence as the organization trained and the individuals learned their roles.

Metric Validation

Figure 16 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.

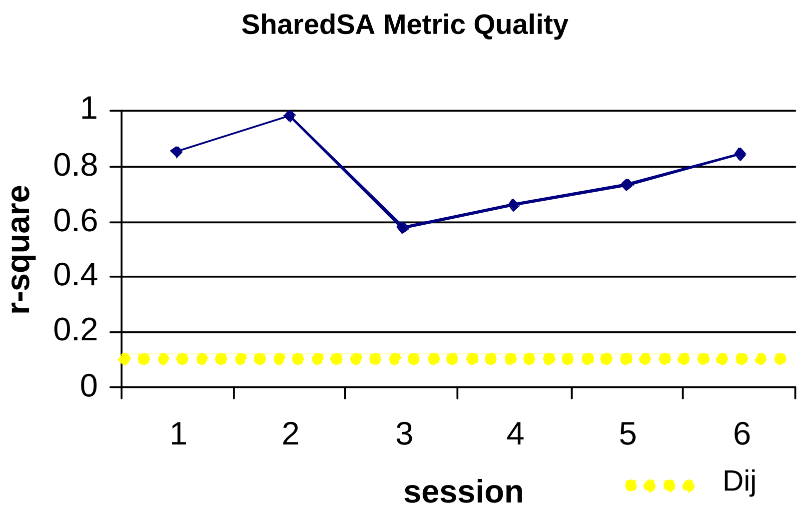


Figure 16: Validation Results of Model 3

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 17 shows how much variance was accounted for by the factors included in the models. 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.

Model 1	$SharedSA_{ij} = \alpha D_{ij}$	$r^2 = .09 - .15$
Model 2	$SituationModel_{ij} = \alpha A_{ij} + \beta P_{ij} + \chi C_{ij}$	$r^2 = .24$
Model 3	$SharedSA_{ij} = \alpha A_{ij} + \beta P_{ij} + \left(\frac{\delta H_{ij}}{\chi C_{ij}} \right)$	$r^2 = .78$

Figure 17. Validation Results of Our Models

6.0 Research Accomplishments

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 using these same factors

Further, the metric is sufficiently valid for application to real-world shared SA tracking in military command and control organizations.

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.