NETSTAR: An Intelligence Analyst's Decision Tool

Elliot E. Entin, Ph.D., Rebecca Grier, Ph.D.,Georgiy Levchuk, Ph.D. (PM), Tyrone Jefferson

Aptima Inc.

Abstract

In order to assess the effectiveness and utility of the NetSTAR system, it is imperative that we have an understanding of current performance levels on the tasks with which it will assist. To that end, we conducted an experiment at NPS with Navy and Marine Officers. The officers were asked to first identify the organizational map that was closest to the data they had and then to map the nodes(commanders, leaders, assets) within that organization. Although the participants performed the organizational identification task significantly better than chance, some organizations proved more difficult to discern than other and as the noisiness of the data increase, as predicted, the accuracy of organizational identification declined.

Introduction

Knowledge of the principles and goals under which an adversary organization operates is critical to accurately predicting an adversary's future activities. To implement successful counter-measures, additional knowledge of the specifics of organizational command, communication, control, and information networks, as well as responsibilities and roles of members comprising the organization, are also necessary. In short, detailed information about the attributes available to the adversary organization and the interrelationships among team members, organizational resources, environment areas, and mission tasks is needed.

Analysis of the behavior of organizations, ranging from the more structured command systems of a conventional military to the decentralized and elusive insurgent and terrorist groups, suggest that a strong relationship exists between the structure, resources, and objectives of those organizations and resulting actions. Thus, when analyzing an enemy organization, we are not merely interested in learning about individuals, but how they are organized as a team and what they can do together. Just like brain structure can be discovered using MRI scans, the structure of the enemy organization can be discovered from observations of interactions and activities of its members; these may be observed as part of normal activity monitoring, or may be induced with intentional probes. In the context of discovering a covert organization, the scope of probes is very limited. Therefore, one needs to use real–life observations obtained as part of normal monitoring, that are tightly coupled with the intent of the adversary.

Organizations perform their missions by accomplishing tasks which may leave detectable events in the information space. The dynamic evolution of these events creates patterns of the potential realization of organizational activities and may be related, linked, and tracked over time (Pattipati *et al.*, 2004). The observed data, however, is very sparse, creating a challenge to connect relatively few enabling events embedded within massive amounts of intelligence data. On one hand, to successfully detect the interaction patterns over time, models are need to associate the observations with true but hidden interactions in the adversarial organization. Such association is at the core of temporal activity pattern identification models (Levchuk and Chopra, 2005) and requires the knowledge of the mapping between the observed actors and the decision-making nodes in the hypothesized enemy organization. On the other hand, this mapping defines the role of the actors and their place in the organization, which is essential to directing the counter-measures against the most important enemy actors or relationships among them.

The meaning of organization discovery, the NetSTAR challenge, is the ability to recognize the command, control, communication, and task structures of an organization. However, the challenge is that most of the time we cannot observe the elements of organizational structure. Instead, we must collect intelligence on the actions and activities of the organization. The specific actions depend on the structure of the enemy's C^2 organization, that in turn are derived from the goals of the specific enemy. The problem of structural network discovery is very complex: the observed data does not relate to the structure directly; instead, it relates to its manifestation in the form of activities and processes that are enabled by the organizational structure(s) and performed by the organization's members. Therefore, whether it be human analysts or algorithms to reconstruct the organization from observations alone both will need to search through a very large space of possible structures. Given historic data and the availability of subject–matter experts, we can reframe the problem as one of hypotheses testing, where a set of predefined hypotheses about the enemy organization and its sub-elements, albeit very large, is given. The problem then becomes one of rank–ordering these predefined hypotheses on the basis of how best they match (or explain) the observed data. This is precisely the problem that faced the human analyst teams in this study.

The graph matching problem has many complicating aspects. First, there exist many mappings from individuals/actors to command nodes (e.g., there are N*M mappings from N actors to M command nodes). Second, we need to explore many different hypotheses about enemy organization – that is, many organizational structures. Third, even if the organization is known, we still need to determine what goals/mission it has, and how far along this organization is in finishing the mission. Other issues, such as transcribing the communications to identify the content, constructing feasible organization and mission representations, and determining the most efficient intervention strategies must also be addressed.

It is hypothesized that identifying such mappings poses a difficult task for most humans even in ideal circumstances. Moreover, in the real world the available/obtainable data to be used for identifying adversarial organization's structures usually consists of partially classified communication transactions among tracked actors (e.g., "members of militant wing engaged in a meeting with weapons suppliers at 11:35 am for 35 min to procure explosives") and their individual actions (e.g., "BLUE team discovered a safehouse and apprehended RED operatives attempting to manufacture weapons"). Such data are very noisy and sparse due to challenges in data collection (e.g. limited sensors and/or human intelligence, security of adversary communication networks, uncertainty in message translation, data association uncertainty, etc). Such "fogging" of the data renders this task even more difficult for humans.

Further, there are several different organization types in the world. More specifically, there are organizations that are structured around functions, where commanders control a certain capability. A second type is organizations that are structured around a multi-functional platform. Finally there are organizations that combine functional and divisional structures. It is within reason, to assume that divisional and hybrid organizations may require more deductive reasoning skills than functional, given the duplicity of capabilities and tasking within commands. Where as in a functional organization, there would be limited replication across commands.

Experiment Overview

The experimental design and methodology described in the next section focused on human-in-the-loop experimentation that investigated the accuracy of organization discovery obtained by a team of human analysts. Later the human based results can be compared to NetSTAR tools to evaluate the value added by the decision tools, however, such a comparison and evaluation are not a part of this paper.

The methodology diagrammed in Fig. 1 leveraged many years of similar model-based experimentation cycles executed for the Adaptive Architectures for Command and Control (A2C2) research program (Diedrich et al., 2003; Entin et al, 2003; Kleinman et al, 2003; and Levchuk et al, 2003). The A2C2 work studied the ability to use models to develop optimized military organizational structures for different missions and to encourage organizational adaptation. The A2C2 program included iterative cycles of experimentation to evaluate and validate the modeling approaches. These experiments have been conducted using the Distributed Dynamic Decisionmaking (DDD) simulation environment (Kleinman, Young, and Higgins, 1996). Successive DDD generations have demonstrated the paradigm's flexibility in reflecting different domains and scenarios to study realistic and complex team decision-making. An outcome of A2C2 program that directly feeds our investigations has been the creation of DDD-based scenarios and organizational structures. The A2C2 experiments have catalogued a diverse set of outcomes from simulation runs for various teams, organizations, and mission conditions.

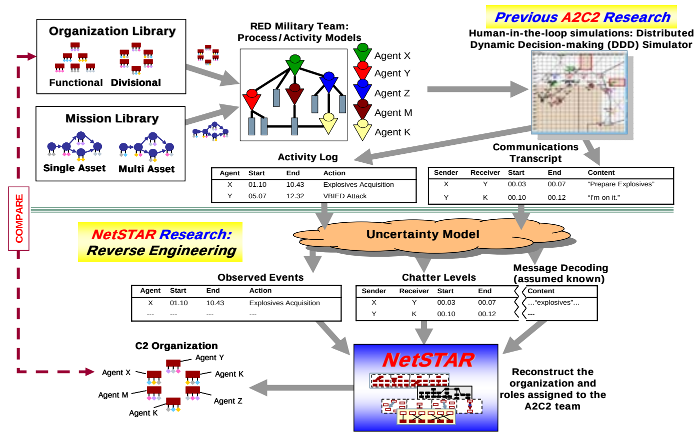

Figure 1: NetSTAR Project Workflow

A human-in-the-loop DDD run includes a team of participants playing roles of commanders in a predefined command and control team and performing the mission tasks in the DDD simulation environment using kinetic and non-kinetic assets/resources. Of particular interest to our investigations are A2C2 experiments with Joint Task Force organizations, which explored the range of possibilities to assign the C^2 relationships, resource ownership, and individual responsibilities among commanders. Under the A2C2 program, we have tested both traditional and non-traditional C^2 structures, thus providing a rich data set to be used as part of the data for the current study. For each human-in-the-loop simulation run from an A2C2 experiment, data logs were captured that include task execution logs (who does what, where, and when) and the communication interactions among team players. The latter information was coded into distinct categories corresponding to several types of formal and informal interactions in a C^2 organization. This data was directly used in our investigations. Specifically, the various data logs from several select trials for functional, divisional, and hybrid organizational structures were directly used to construct the hypothesis organizational structures used in the study. The log files were are also the source of the input data sets that the human analysts were provided – its these data sets the analysts is attempting to match to the hypothesis organizational structures to discern the organizational structure of the input data sets. However, the input data sets were not noiseless. An uncertainty model was used to insert noise – that is, introduce deceptive events (false alarms), create missing data (misdetection), and errors to the input data.

The experiment investigated two specific questions:

(1) Are some types of organizational structures more difficult for human analyst to discern?

(2) How much noise in the data can human analyst deal with and still come up with solutions that are superior to chance?

To address the first question, we manipulated the type of organization over three levels. The A2C2 research provided three distinct organizational structures: functional, divisional, and hybrid which is essentially a blend of functional and divisional. In a functional organization each commander controls and can be identified with a single type of asset, e.g., all air assets or all surface assets. Where as, in a divisional structure each commander control a platform and the myriad of assets on that platform. We hypothesize that it may be easier to discern a functional than a divisional organizational structure. When hybrid organizational structures are added to the mix that have both functional and divisional qualities they should prove to be the most difficult to recognize. All this assumes noiseless data on which to base identification decisions – an unlike situation.

Thus, the second question addresses the impact of noise on the identification process. The prudent hypothesis is to assume that the higher the noise level the more difficult the identification process. But it may turn out that some organizational structures are more or less affected by noise. We hypothesize that given their mixed nature the hybrid organizational structure will be most difficult to discern in noisy data environments.

Method

Participants

Nine two-person teams drawn from a class at the Naval Postgraduate School (NPS) were included in this study. The sample was comprised of four Marine officers, thirteen Naval officers, and one civilian. Participants' years of military service ranged from 1 to 21years.

Independent Variables and Design

Two independent variables were manipulated: organization type and fogging level. Organization type was operationalized as varying organizational structures along a continuum, ranging from functional to divisional organizations. Functional organizations were organized such that each commander specialized in one or two aspects of a mission such as Strike or Air Warfare, where the specific assets controlled were distributed across multiple platforms (ships). In contrast, divisional organizations were designed so that each commander had control over a single multifunctional platform that was capable of processing a variety of tasks in a given location. There was an intermediate, or hybrid organization designed so that of the six-person team four commanders controlled assets divisionally and two controlled assets functionally.

The second independent variable, fogging level, referred to the amount of noise or error present in the data tables and illustrations describing an organizational structure. Noise and error were introduced through the random removal of different messages or nodal connections, the inclusion of false or deceptive messages, or mislabeled messages. Using a specific algorithm three levels of fogging were produced (10% fogging, 30% fogging, and 50% fogging). Specifically the information and data describing five stimuli organizations were corrupted to produce descriptive information sets for five organizations: Functional with 30% fogging, Divisional with 30% fogging, Hybrid with 10% fogging, Hybrid with 30% fogging, and Hybrid with 50% fogging.

Although various constraints precluded complete crossing of organizational structure and fogging, it was possible to compare the three organizations at the same level of fogging and to compare the three levels of fogging within one organization type. All teams saw the functional, divisional, and two of three fogged hybrid organizations. Each of the three fogged hybrid organizations was presented an equal number of times across participants. Four teams saw the divisional organization in the first trial and four teams saw the functional organizations in the first trial. The final team saw the hybrid organization with a 30% fogging level in the first trial. The presentation of the conditions was counterbalanced in the remaining three trials.

Materials

Data sets for seven different hypothesis organizational structures were prepared and presented to the participants across all trials. A hypothesis organizational structure described an organizational structure that might match the stimulus structure. The seven hypothesis organizational structures ran the continuum from functional to divisional structures (i.e., one functional, one divisional, and five intermediate or hybrid structures). Descriptions of the organizational structures were presented in both spreadsheets and diagrams that, in turn, described the interaction among nodes comprising the organization. Nodes represented the items of interest in the organizations including commanders, leaders, assets, and areas. No noise or errors were introduced to the hypothesis organizational structures. Each hypothesis organization was comprised of nine descriptive spread sheets and nine corresponding diagrams. To help the participant teams keep the seven hypotheses organizational structures distinct, data for each hypothesis organizational structure was printed on a different color paper and placed in its own folder.

The stimulus data (i.e., the data sets to which participant teams were trying to match to the hypothesis organizational structures) also consisted of the same nine descriptive spread sheets and nine corresponding diagrams as in the hypothesis sets.

Dependent Measures

The primary dependent measures assessed how accurately the participant teams were able to identify the organizational structure of the noisy stimulus organization they were given by selecting the correct hypothesis organizational structure that matched it, how accurately they could identify the command nodes of the stimulus organization, and how accurately they could identify the platform nodes of the stimulus organization. In addition several self-report measures were assessed to gain insight into the participants' cognitive processes as they perform the tasks.

The number of times a stimulus organization was correctly matched to a hypothesis organizational structure was counted across all teams and evaluated against chance. To derive a selection accuracy score participant teams were also asked to indicate on a prepared form the percent match of the stimulus data set to each of the seven hypothesized organizational structures. The selection accuracy score measured how closely the chosen hypothesis organizational structure was to the correct one. To do this, a matrix was computed that indicated the percent overlap between any hypothesis organizational structure to any other. This was done in the same manner described by Weil et al (2005). Thus, if a participant indicated the correct hypothesis organizational structure a score of 100 was given. If the indicated hypothesis organizational structure was not the correct structure, but overlapped 83% with the correct one, a score of 83 was given.

To assess how accurately participant teams mapped the correct command onto the commander nodes of the stimulus organization teams were asked to indicate on a prepared form the commander's name in the stimulus data set that was associated with the commander's name in the chosen hypothesis organizational structure's data set. A percent agreement score was computed. That is, the number of correct associations made was divided by the total possible, in this case six, to derive the percent agreement score. In a similar manner the correctness of mapping the leader to correct platform node was also assessed and a percent agreement score computed for the leader to platform mapping.

A self-report measure of workload was obtained from the participants using the Task Load Index (TLX; Hart & Staveland, 1988). The measure asked participants to indicate on a scale from 1- 20 the level of Mental Demand, Temporal Demand, Effort, Frustration, and Performance experienced on the task just completed. These scores were then averaged to provide a workload score for each participant at the end of each trial. Participants were also asked to self-report their confidence in their task solution at the end of each trial on a scale from Not at all Confident (1) to Completely Confident (6), the amount of fogging that was present in each stimulus organization on a scale from 0 to 100, and perceived complexity of the task on a scale from Not at All Complex (1) to Completely Complex (6).

Procedure

Four days prior to the first trial, all participants received an hour long briefing on the nature of the experiment and were provided with training for the experiment. The training consisted of describing to the participants the nature of the problem they would be asked to solve, a general description of the data they would be examining, and strategies for approaching this form of problem. At the end of the training session participants were asked to sign an informed consent form.

Teams came to the lab for 2 two-hour sessions. During each one hour trial, participants were provided one stimulus organization data set and 7 hypothesis organizational structures. Participants had 50 minutes to select the correct hypothesis organizational structure and get as far in the commander, leader, and asset mapping as possible. At the start of the trial, participants were told they had 45 minutes to select an organization. If they had not indicated a hypothesis organization by that time, they were asked to make their best guess and to begin filling out the mappings. Throughout the trial the amount of time left was indicated. After teams completed the first trial materials were gathered and participants were then provided a different stimulus organizational data set and the process was repeated. During most sessions two or three teams were working in the lab simultaneously.

Results

Measures Evaluating Selection and Mapping Accuracy

The percentage of teams who selected the correct hypothesis organization is presented for each condition in Figure 2. Overall the number of correct solutions was 17 or 47.2%. On average each team selected the correct hypothesis structures on two of the four trials. If only chance were operating a team would have had only a one in six chance (16.7%) of selecting the correct hypothesis organization on any given trial. We can see that participant teams performed considerably better than chance ($Z = 4.91$, $p < .01$). We can also see from Fig. 2 that treatments varied in difficulty, only 1 out of the 6 teams that received the 50% fogged data chose the correct hypothesis organization. This suggests that when the stimulus data contained a great deal of errors participants had a difficult time identifying the correct structure. In fact they did no better than chance.

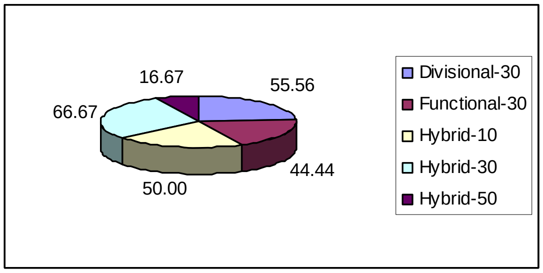

Figure 1. Percentage of Teams Who Selected the Correct Hypothesis Organization by Organizational Type

We now turn to the analysis of selection accuracy in terms of percentage over lap between the selected hypothesis organization and the correct one. The mean selection accuracy for all the experimental conditions is depicted in Fig 3. Recall that an omnibus within-subjects analysis across all conditions is not possible because not all teams saw the same set of experimental conditions. Instead each hybrid condition was contrasted one at a time with the Divisional-30 and Functional-30 conditions. These analyses revealed that (1.) Functional-30 condition was performed less accurately than the Hybrid-10% conditions ($p < .05$), (2.) the Hybrid-50% condition was performed less accurately than the Divisional-30 condition ($p < .05$), (3.) the Hybrid-50 condition was performed marginally less accurately than the Functional-30 condition ($p \leq .08$). Three hybrid conditions were contrasted and the analyses showed that the Hybrid-50% condition was performed less accurately than the Hybrid-30% and Hybrid-10% conditions ($p < .05$). For this latter result it would appear that increased fogging resulted in reduced accuracy. No other contrasts for the selection accuracy score proved significant.

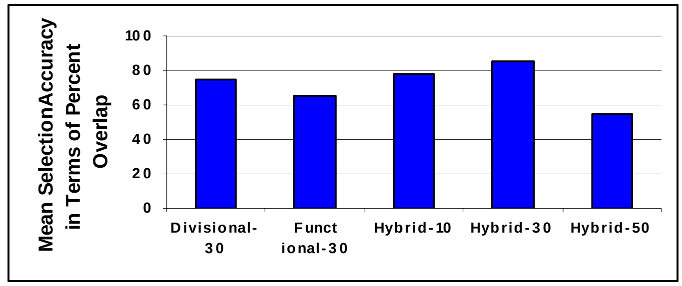

Figure 3. Mean Selection Accuracy in terms of Percent Overlap by Organizational Type

The percent correct commander mapping means for all organizational structures are shown in Fig. 4. As we can see participant teams attained a higher percent correct mapping commander nodes in the functional organizational structure than for any of the other organizational structures. More specifically a statistical difference was observed when the Functional-30 condition was contrasted with the Divisional-30 and Hybrid-10 conditions ($p < .05$) and contrasted with the Divisional-30 and Hybrid-50 conditions ($p < .05$). When we pooled the three hybrid organizational structures and analyzed the pooled hybrid condition with Divisional-30 and Functional-30, the Functional organizational structure showed significantly higher correct commander mappings ($p < .05$).

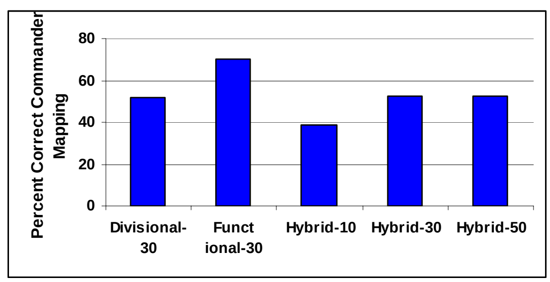

Figure 4. Mean Accuracy of Commander Mapping by Organizational Type.

Analyses of the means for the percent correct platform mapping revealed no significant results. It should be noted that platform mapping was completed on only 60% of the trials compared to almost 100% for commander mapping.

Analysis of the Self-Report Data

Examining the overall means for the self-report measures depicted in Table 1, we can see that the means for complexity and confidence exceed 4.0, the "slightly confident" and "slightly complex" categories. A one mean t-test shows that both means significantly exceed 4.0 ($ps < .01$) and a check of the frequencies indicate that 60% of the participants exceeded a rating of 4.0 for complexity and 69% did so for confidence. Participants are therefore reporting that overall the tasks were moderately to highly complex and that they were moderately to highly confident in their solutions.

Table 1. Overall Means for End of Trial Measures

Measure	Mean	Std Dev	N
Complexity	4.90	0.80	71
Confidence	4.47	1.30	72
Fog Perception	46.25	19.48	72
Fog Difference	20.56	17.89	72
Workload	65.01	12.77	72

Participants' perception of the amount of fogging they were encountering on the different trials did not agree with the actual degree of fogging. The computed correlation between the estimated fogging level and the actual was $r = 0.08$, and was not significantly different from zero. Kendall's coefficient of concordance W equaled 0.26, again showing very little agreement between the actual fogging amount and what participants estimated. Apparently, participant experienced a difficulty with gauging the amount of noise (fogging) present in the information they had to deal with.

To further examine the self-report measures the five measures were correlated and the resulting intercorrelation matrix is presented in Table 2. Not surprising the higher the perceived complexity of a problem the higher the reported workload and the lower the reported confidence. Moreover, as perceived workload goes up reported confidence in one's solution goes down and higher perceived fogging brings about lower confidence.

Table 2. Inter-correlation of End of Trial Measures

	WKLD	FogPerc	FogDiff	Cplex	Confid
Workload	1				
FogPercept	0.07	1.00			
FogDiff	0.00	**0.85**	1.00		
Complex	**0.46**	0.19	0.03	1.00	
Confid	**-0.54**	**-0.44**	**-0.28**	**-0.37**	1.00

Note- Coefficients in boldface significant at $p < .01$, one tail.

To test the independence of the confidence ratings between conditions we recoded the confidence ratings about the median in to low (ratings 1 through 4) and high (ratings 5 through 6) and contrasted the conditions in bivariate

tables. Confidence ratings in the functional and divisional conditions were significantly related (Fisher's Exact Probability Test, $p < 0.02$), high confidence ratings in the solution for the functional conditions were related to high ratings in the divisional conditions and low confidence ratings in the functional conditions were related to low ratings in the divisional conditions. None of the other analyses produced significant results.

The means for perceived fogging are presented in Table 3. As we can see from this table participants tended to overestimate the amount of fogging present in all conditions except in the Hybrid-50 condition. The analyses comparing the condition means revealed that the Hybrid-30 was significantly higher than the Divisional-30 and Functional-30 conditions ($p < .05$) in the Hybrid-30 Block (i.e., the 12 participants that experienced these three conditions). No other reliable differences were found.

Table 3. Mean Perceived Fogging for the Five Conditions.

	Division-30	Funct-30	Hybrid-10	Hybrid-30	Hybrid-50
Mean	41.11	45	45.42	52.50	50.42
SD	14.81	21.63	20.50	25.89	13.73
N	18	18	12	12	12

The means for the error in fogging estimates are depicted in Table 4. Analyses show that participants made the greatest errors estimating the fogging in the Hybrid-10 condition ($p < .05$) and lowest errors estimating the fogging in the Hybrid-50 condition ($p < .05$). Error was also marginally larger in the Functional-30 than Divisional-30 condition ($p < .06$). It would appear that when the true level of fogging was low
participants made the largest error when estimating the fogging. The reverse appears to be true when the actual fogging is high, participants made the smallest error in estimating fogging. The functional organizational structure seemed to engender larger error than the divisional organizational structure.

Table 4. Mean Error in Fogging Estimates for the Five Conditions

	Division-30	Funct-30	Hybrid-10	Hybrid-30	Hybrid-50
Mean	13.89	20.56	35.42	26.67	9.58
SD	12.07	16.08	20.50	21.14	9.40
N	18	18	12	12	12

Discussion and Conclusion

The primary goal of this study was to evaluate human analysts' performance in the identification of organizational structures, particularly at varing levels of data accuracy. This study showed that humans are capable of identifying organizational structures at a level well above chance and particularly when the data were relatively error-free. However, as predicted, identification performance degrades as the data became noisier. There was also evidence that some organizational structures were more difficult to identify. Specifically, the it was the functional organizational structure that appeared to present the most difficult for participants to discern. Once teams faced the task of mapping commanders to nodes or leaders to platforms the functional organizational structure appeared easier than the other two structures. There are several empirical reasons for these results. The overlap, i.e., the closeness, among the hypotheses organizational structures is not uniform from one to another. As shown in Figure 4, the overlap between the F2 structure and the F structure is 83% - not far apart – and the overlap between F3 and F is 70%. When we make the same comparisons for the hypothesized divisional organizational structures the picture is quite different. D2 only overlaps 51% with D and the overlap with D3 is about the same. Thus, it was empirically more difficult to discriminate between F and F2 or F3 than it was to discriminate between D and D2 or D3. After the hypothesis organizational structure was selected and the participant team set upon the task of mapping commanders to nodes and leaders to platforms, the discrimination problem "flips." Under a functional organizational structure a commander owns only one type asset (e.g., all anti-air assets or all surface assets), while under the divisional organizational structures commanders own a little of each type of assets. Thus, the discrimination task under the functional structure is easier than under the divisional structure.

In addition to gaining a better understanding of performance, several insights into human cognition were achieved through the self report measures. First, participants were able to accurately recognize their ability to perform the task. On average participants reported that the task was slightly complicated. Further on average they reported that they were slightly to moderately confident with their performance. Given that participants did indeed perform better than chance but not perfectly, the confidence ratings appear to be an accurate representation of their abilities to perform this task in general.

Second, participants generally overestimated the amount of fogging in the data. This impacted their ability to identify the proper solutions. A frequent comment when the participants were solving the problem when a proposed solution didn't match identically with the organization they were viewing was that it was probably noise in the data. That is, the participants were affected by the confirmation bias (Kahneman and Tversky, 1979), which states that people look for information that supports a proposed solution and ignore data that refutes a hypothesis. Overestimating the amount of noise in the data is a subconscious way to approve a hypothesis that does not necessarily match the data provided.

In conclusion, the study demonstrated that human decision makers are capable of working with "noisy" observed data and discerning from a set of hypothesized organizational structures the organizational structure that produced the observed data, and to do so well above chance. We also observed that as participants attempted to map more detailed aspects of the organizational structure their ability sharply decreased with each increased levels of detail. This is most likely due to the fact that humans cannot consider all the information required to make these more detailed mappings. Computer algorithms like NetSTAR do not suffer from such constraints and hold the potential to greatly aid in the task of organization identification and description.

References

Diedrich, F.J., et al., "When do Organizations Need to Change (Part I): Coping with Incongruence", *Proceedings of the 2003 International Command and Control Research and Technology Symposium*, Washington, DC, 2003

Entin, E. E., et al., "Inducing Adaptation in Organizations: Concept and Experiment Design", *Proceedings of the 2004 Command and Control Research and Technology Symposium*, San Diego, CA, 2004

Entin, E. E., et al., "When do Organizations Need to Change (Part II): Incongruence in Action", *Proceedings of the 2003 International Command and Control Research and Technology Symposium*, Washington, DC, 2003

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.), *Human mental workload*. Amsterdam: Elsevier.

Kahneman, D. and Tversky, A (1979),. 'Prospect Theory: An Analysis of Decision Under Risk'. *Econometrica*, 47 263-291.

Kleinman, D.L., et al., "Scenario Design for the Empirical Testing of Organizational Congruence", *Proceedings of the 2003 International Command and Control Research and Technology Symposium*, Washington, DC, 2003

Kleinman, D.L., P. Young, and G.S. Higgins, "The DDD-III: A Tool for Empirical Research in Adaptive Organizations", *Proceedings of the 1996 Command and Control Research and Technology Symposium*, Monterey, CA, June, 1996

Levchuk, G.M., Chopra, K. (2005). "NetSTAR: Identification of Network Structure, Tasks, Activities, and Roles from Communications", *Proceedings of the 10th International Command and Control Research and Technology Symposium*, McLean, VA.

Levchuk, G.M., et al., "Congruence of Human Organizations and Missions: Theory versus Data", *Proceedings of the 2003 International Command and Control Research and Technology*

Symposium, Washington, DC, June, 2003

Pattipati, K., Willett, P., Allanach, J., Tu, H., and Singh, S. (2004). “Hidden Markov Models and Bayesian Networks for Counter-Terrorism”, in *21st Century Enabling Technologies and Policies for Counter-Terrorism*, Robert Popp and John Yen eds. (in press)

Weil, S.A., Levchuk, G., Downs-Martin, S., Diedrich, F.J., Entin, E.E., See, K.E., and Serfaty, D. (2005). Supporting organizational change in command and control: Approaches and metrics, *Proceedings of the 2005 Command and Control Research and Technology Symposium*, Washington, D.C.