19th ICCRTS

Paper 038

Title: Human Limits to Cognitive Information Fusion in a Military Decision-Making Task.

Topics: Topic 4: Experimentation, Metrics, and Analysis; Topic 3: Data, Information, and Knowledge; Topic 5: Modeling and Simulation

Authors: Laura R. Marusich, Norbou Buchler, Jonathan Z. Bakdash

Point of Contact:

Laura R. Marusich
U.S. Army Research Laboratory, Human Research and Engineering Directorate
Aberdeen Proving Ground, MD 21005
254-288-9377
laura.r.marusich.ctr@us.army.mil

Abstract:

One purported benefit of Network Enabled Operations is the increased availability of shared information, which can contribute to improved situational awareness, decision-making and overall mission effectiveness. Using a simulated Mission Command task focused on capturing high value targets (HVTs), we investigated how varying levels of available information affects human decision-making. The information presented to participants consisted of reports of possible HVT locations. Some reports indicated the correct location of a target; incorrect reports indicated a variable location immediately adjacent. As compared to a single report, it was possible to pinpoint the true location of a target by integrating information from multiple reports. However, results showed that participants did not locate HVTs faster with multiple intelligence reports compared to a single report. To determine if this was due to cognitive limits, human performance was compared to an Ideal Observer Model which had perfect information integration but the same task timing constraints. The model demonstrated a considerable improvement in performance with increased volume of information. These findings raise questions about human capabilities for information fusion given the high volume of information in military networks. Furthermore, results suggest decision support systems may enhance human capabilities for fusing and disambiguating information.

Human Limits to Cognitive Information Fusion in a Military Decision-Making Task

Introduction

Recent advances in technology have transformed the way that people communicate and organizations operate in society today. This is perhaps most evident in the military domain, where these advances have permitted a dramatic change in the way that command and control (C2) networks function. The military's transformation to Network Enabled Operations (NEO) gives rise to large, interacting, and layered networks of Mission Command personnel communicating and sharing information within and across various command echelons (Alberts & Garstka, 2004; CNSFAA, 2005). With such a dramatic shift to extensive networking, there is a need to understand the impact of NEO at the level of human cognition, in particular human decision-making in such complex, information-rich environments. In time-stressed situations common to the Mission Command environment, the performance of the entire networked organization can be constrained by the ability of a single Soldier's ability to process information in a timely manner.

The transformation of the U.S. and NATO countries to NEO has proceeded under a conceptual framework of Network Centric Warfare comprising four primary tenets (Alberts & Garstka, 2004):

1) A robustly networked force improves information sharing and collaboration.
2) Such sharing and collaboration enhance the quality of information and shared situational awareness.
3) This enhancement, in turn, enables further self-synchronization and improves the sustainability and speed of command.
4) The combination dramatically increases mission effectiveness.

This conceptual framework explicitly assumes that greater information sharing in a networked organization produces better situational awareness, decision-making, and ultimately, mission outcomes. In essence, the increase in information available to commanders and their staff is postulated to increase the quality of decision-making due to enhanced situational awareness (CNSFAA, 2005).

The tenets of this framework have not been investigated at the level of human cognition, especially in relation to decision-making and human information processing. For instance, there may be situations in NEO where increased information sharing raises the *quantity* of available information without a corresponding increase in *quality*. This presents a challenge as cognitive resources must be devoted to separating the relevant information (the signal) from the irrelevant information (the noise). Even when information sharing results in the increased availability of relevant information, the sheer volume and rapid pace of information received and readily accessible through networked systems can be overwhelming. Humans have a fixed cognitive processing capacity limited by attention, memory, as well as the availability and communication of information. In complex information environments, it can be increasingly difficult to pinpoint and fuse the relevant information to support decision-making.

Understanding how humans process information and make decisions in relation to information flow is a critical challenge that we wish to address through experimentation. How much information is too much and what happens to decision-making when an operator becomes overwhelmed? Understanding of the consequences to human performance of operating in an information-rich, time-stressed environment is a critical challenge that fits within the Office of the Secretary of Defense 'Data to Decisions' initiative to manage the complexity of the information environment in ways that "enable faster, better decisions while reducing information overload" (Swan and Hennig, 2012).

Consequently, for NEO there is a need to examine human cognition, specifically objective decision-making in controlled experiments and observational studies. In this work, we experimentally manipulate the volume of information to test human performance for making decisions to find HVTs using a simulated task. In our case, *volume of information* is also related to the concept of information overload, which has many definitions, including more relevant information than an individual can process or being inundated with irrelevant and/or unrequested information (see Edmunds & Morris, 2000).

Here we focus on volume of relevant information. Thus, the cognitive task of interest is information fusion rather than filtering. We use the term cognitive information fusion to indicate the involvement of humans and to differentiate this from other types of information fusion (Blasch, Bosse, & Lambert, 2012). It is clear that large amounts of bad or redundant information are undesirable. However, there is a technological goal and intuitive desire is to provide the Soldier with as much "good" information as possible so that decisions can be made based on the most complete understanding of a given situation possible. But what if too much "good" information can also impede performance? By studying cognitive information fusion, rather than filtering, we can address exactly this question.

Cognitive Information Fusion and Information Overload

Contrary to intuition, research shows that more information (even highly-relevant information) does not necessarily lead to better decision-making. Nadav-Greenberg and Joslyn (2009) asked participants to make repeated decisions as to whether or not to salt the roads in a town, based on their prediction of whether or not it would freeze on a given night. They were given the expected overnight temperature, and the between-subjects manipulation was what additional information the participants received: no additional information, the lower bound of the 80% confidence interval on expected temperatures, the lower and upper bound of the CI, the probability of freezing, or the option to request any or all of these pieces of information. Participants in this last condition, with all types of information available to them upon request, actually performed worse than those in the other information conditions. This outcome provides some evidence for information overload; however, some of the types of information provided were redundant. For example, the probability of freezing was calculated from the 80% confidence interval. Our interest is in studying the impact of providing more useful information rather than the impact of presenting the same information in a variety of ways.

In another study, participants predicted whether a firm would experience financial distress within three years on the basis of 4, 6, or 8 different information cues (Chewning & Harrell, 1990). The researchers found that approximately one third of their participants demonstrated a u-shaped relationship between information used and information available; i.e., they showed signs of information overload. Those participants who demonstrated overload made less consistent decisions. This study is relevant to our work in that the information provided was not redundant and was varied systematically. However, there was no ground truth with which to compare the participants' decisions. Our goal was to design a study that directly related volume of information to decision-making performance.

For our investigation, we manipulate the amount of information (all useful) presented to the user in performing a task. Essentially, our work examines human capabilities for cognitive information fusion. Information fusion is defined as the integration and merging of information from heterogeneous sources. Cognitive information fusion refers to the role of a human user in integrating information into a conceptual mental model or representation (see Blasch et al., 2012). In designing this study, we have two divergent hypotheses:

1) **"More is More:"** More task-relevant information leads to better performance. This hypothesis is supported by the tenets of NEO – more information sharing leads to greater SA and mission effectiveness.
2) **"More is Less:**" More task-relevant information leads to worse performance. This hypothesis is drawn from the information overload literature, referenced above.

Each hypothesis can also be similarly interpreted by its complete opposite for information and performance. Another interpretation for the first hypothesis is "Less is Less," where less task-relevant information leads to worse performance. An alternative interpretation for the second hypothesis is "Less is More," where less task relevant information results in better performance.

Simulated C2 Task

We created a task designed to simulate a simplified C2 mission. In this task, the participant viewed a computer display containing a grid-based map, as well as several controls and text boxes for the display of information. The participant had control over the movements of four identical assets or units, which could be assigned to travel to any location on the map. The assets took time to travel across the grid. The primary goal for the participant was to find and capture high value targets (HVTs), which activated somewhere in the area of operations (AO) at specified times and remained in the same location until captured. Participants received information about the possible location(s) of HVTs through text updates on the display.

While this task was an obvious simplification of a real C2 scenario, the experimental design conferred the benefits of controlled manipulation of relevant factors. Therefore, strong causal inferences could be made about volume of information and human performance. For example, we systematically manipulated information volume levels by varying the number of intelligence updates presented to the participant for each target. In the real world, information volume is likely to be confounded with the quality/relevance of information, number of sources of information, information modality, rank, echelon, network bandwidth, system availability and interoperability, security restrictions, and many other factors. By using an abstracted experimental task, we were able to hold such potentially confounding variables constant to explore the effects of varied information volume.

Another benefit of this controlled experimental task was that it allowed for straightforward, direct measurement of task performance. In a real scenario, performance is notoriously difficult to measure, whether quantitatively or qualitatively. Even the most high-level measure, mission success, is often ambiguous. In this task, however, we operationalized task performance as the time to capture each target as performance. The degree of success in interpretation and integration of intelligence information determined how quickly participants were able to move assets to the correct target locations. Thus, time-to-capture served as a useful quantitative measure of task performance.

Ideal Observer Model

We developed a model with which to compare human performance data, based upon the concept of ideal observer analysis, originated in the field of perception. The purpose of an ideal observer is "to determine the optimal performance in a task, given the physical properties of the environment and stimuli" (Geisler, 2006, p. 825). Our Ideal Observer Model is an information fusion algorithm that performs the simulated C2 task by integrating all of the information presented to the user. The algorithm receives the same intelligence updates in the same sequence and with the same timing as the human participants. After the first intelligence update, the algorithm assigns the closest unit to the grid location specified in that update. After each new update, it uses the information provided in previous updates as well as the specified location probabilities of the task (see Figures 2 and 3) to generate an optimal prediction of the target's most likely location. In some cases, multiple updates provide enough information for certain knowledge of the target location. In other cases more than one location may be equally likely; in these instances the algorithm makes a random "guess" for its prediction. If a unit is en route to one location, and subsequent updates have confirmed with certainty that the target is in a different cell, the model will stop and reassign the unit. If a unit arrives at a predicted location and doesn't capture a target there, the algorithm updates its list of possible target locations and tries each remaining possibility in turn until the target is captured.

This model is useful in that it defines performance for an ideal observer against which we can compare human performance. We can see, given the context of this particular task, what perfect information fusion would look like in the data, and compare this to actual human performance data.

Method

Participants

Twenty-four volunteers (16 male, 8 female) completed this study. All were between the ages of 18 and 60 years. Participants were recruited through email solicitation at the U.S. Army Research Laboratory, and they did not receive compensation for their participation.

Task

The computer-based task presented to the participants was to find and capture HVTs within a given AO. The display showed a grid map of the AO (see Figure 1) with blue icons indicating the locations of four controllable units available for assignment.

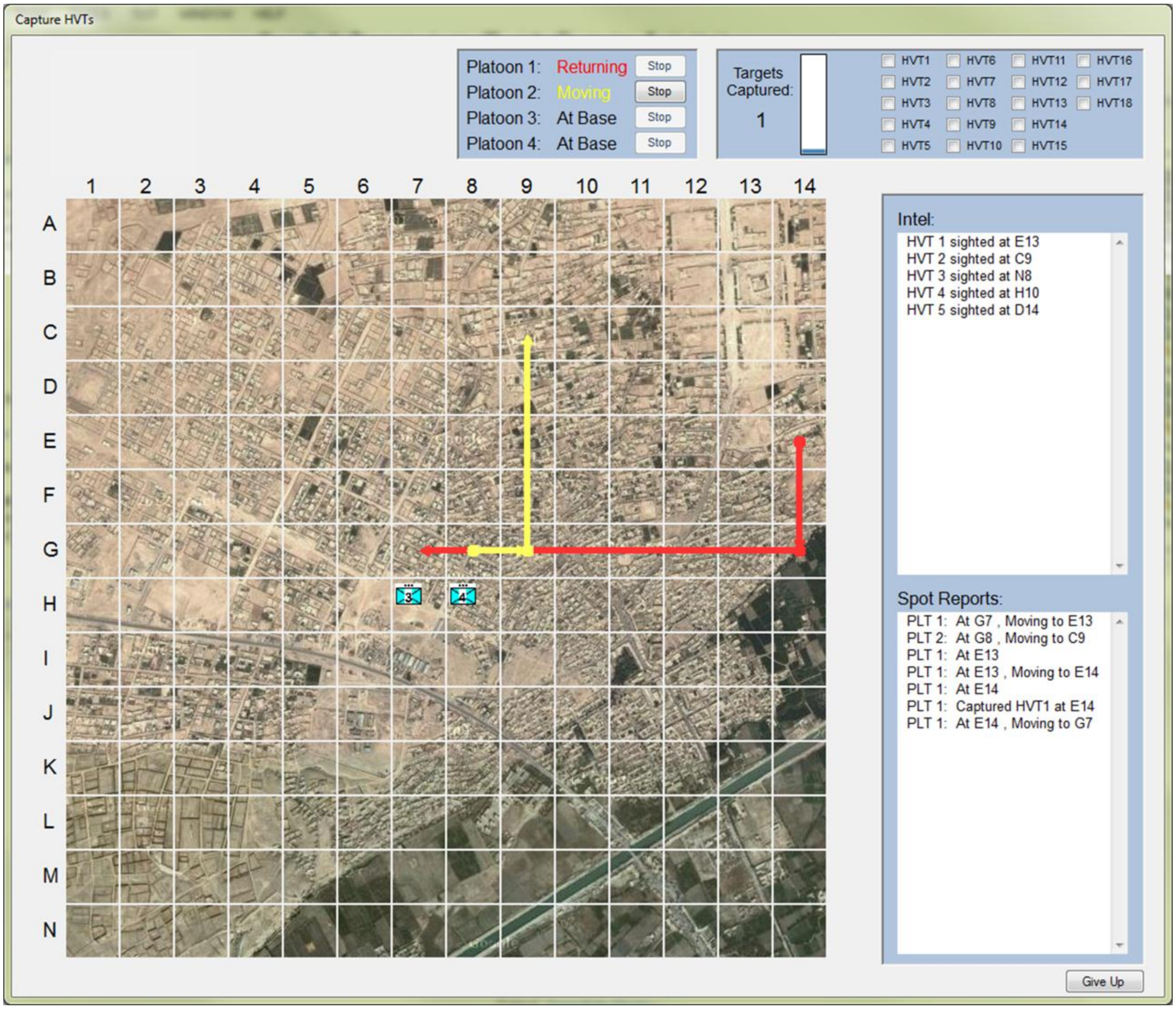

Figure 1. Screen capture of the experimental display.

A “base location” was defined by four grid squares at the center of the map. The four controllable units were located in this base location at the beginning of each round of game play, and they automatically returned to this location after capturing a target. The display also contained a text box that displayed incoming intelligence information about the location of HVTs, a running text box of spot reports issued from the four units, and a progress bar indicating how many HVTs had been captured. Clickable checkboxes were also displayed for each target; participants could use this feature to mark and keep track of which HVTs had been captured and which were still "at large."

As intelligence information appeared, the participant was able to click on unit icons to assign them to travel to these locations and capture the targets. While a unit was traveling, the unit icon disappeared, and a yellow arrow appeared showing the path of travel. Units traveled by taxicab distances, and whether they traveled horizontally or vertically first was randomly assigned for each unit movement. If a unit entered the same location as an active target, it automatically captured the target and returned to the base location in the center of the map. When this occurred, a red arrow appeared showing the path of travel back to the base location. Units always traveled one block every three seconds.

The intelligence updates presented to the participants were 50% likely to provide the accurate location of a target. If the update was not accurate, it was only off by one square in the horizontal or vertical direction (see Figure 2).

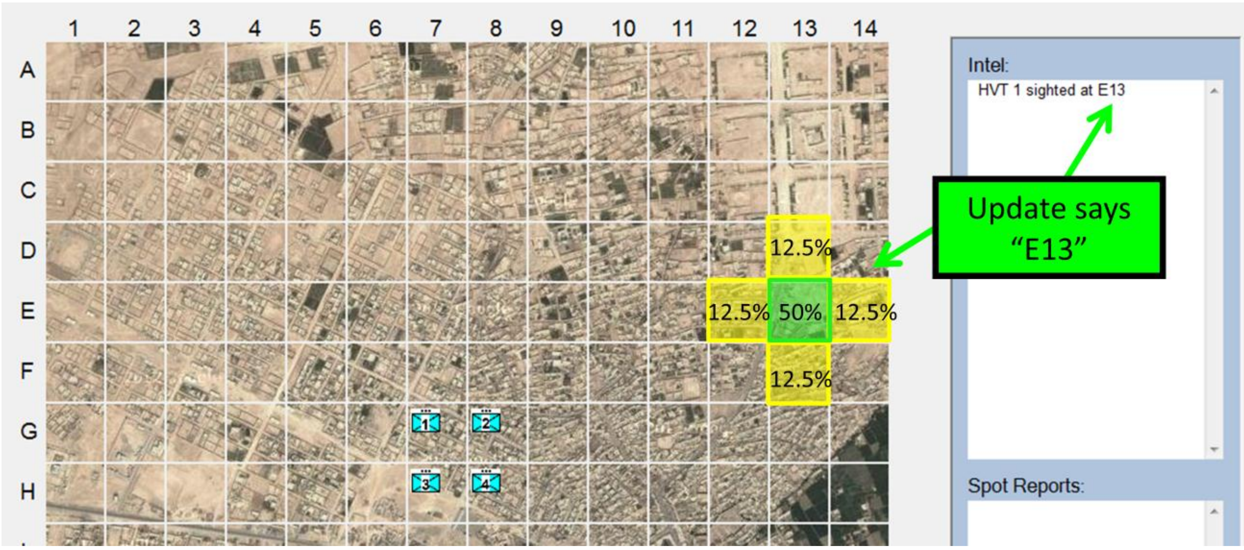

Figure 2. Illustration of possible HVT locations and their respective probabilities, given a single intelligence update. This illustration was shown to participants during the tutorial phase, but it was not part of the experimental display.

As a result of these contingencies, multiple intelligence updates allowed for the possibility of pinpointing the actual location of a given target with certainty. For example, if one update read "HVT 1 sighted at C7," and another read "HVT 1 sighted at C9," the target must have been at C8 (see Figure 3).

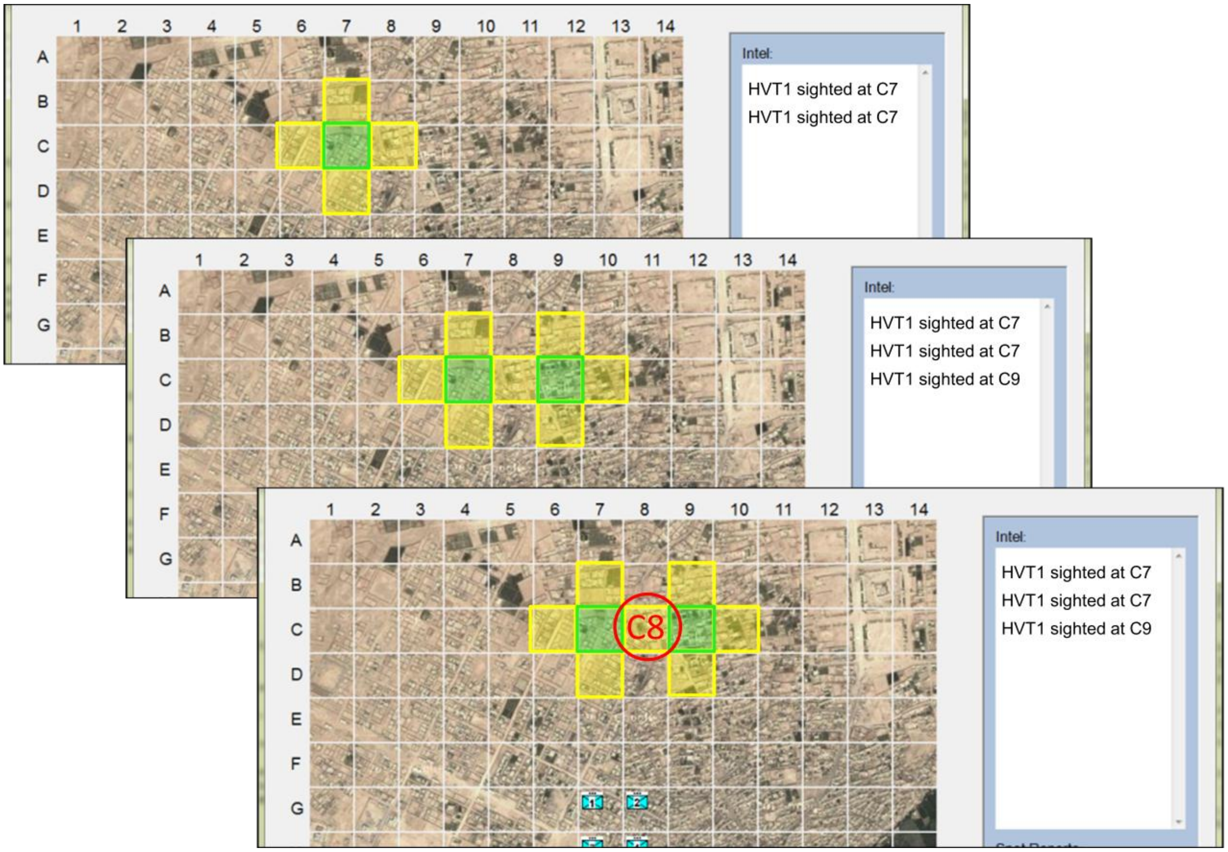

Figure 3. Illustration of integrating multiple intelligence updates to determine the true target location. Given the possible target locations associated with each of the two unique updates, the only possibility is C8. These illustrations were not part of the experimental display.

Participants would receive 1, 5, or 9 intelligence updates per target in a single block. All updates about a single target would appear within a 16-second window, and new targets activated every 15 seconds (see Figure 4).

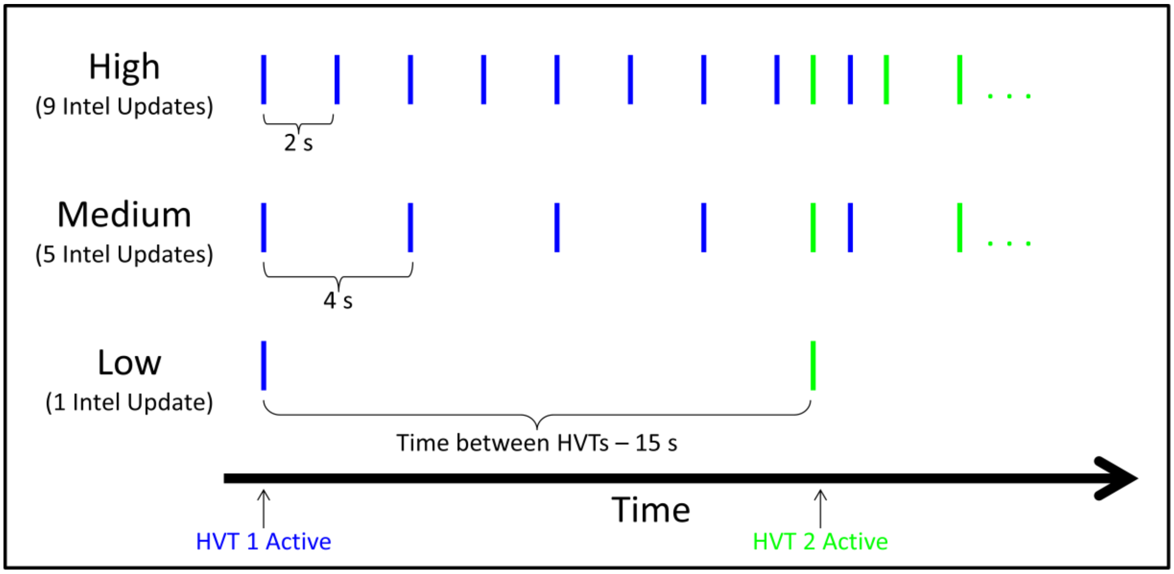

Figure 4. Illustration of the timing of intelligence updates in each of the three information volume conditions.

Procedure

The study was conducted in a sound-attenuated room with a single-monitor computer. Participants completed a self-paced tutorial which provided an overview of the purpose of the task and allowed them to step through each of the actions required of them in the task (reading intelligence updates, assigning a unit to a new location, marking a checkbox to indicate target capture). The accuracy contingencies of the intelligence updates were also described in the tutorial, with a diagram to explain (see Figure 3). Participants then completed a practice block, in which they had to capture six HVTs. Each of the three intelligence volume conditions was presented twice in this practice block. After successfully completing the practice, participants completed three test blocks. Each test block consisted of 18 HVTs to capture. The volume of intelligence updates was varied by block, and the order of the three blocks counterbalanced across participants.

Analysis and Results

For each participant, the time between activation of an HVT (the time of the first intelligence update) and the capture of that HVT was calculated. This time was divided by the distance in blocks of the HVT's location to the base location, to account for the longer travel time required by farther away targets, generating a rate (time to capture/distance traveled). Because there was a great deal of variability in overall speed across participants, rates were converted to standardized values (z-scores) for each participant. Average z-score rates were calculated in the three information volume conditions for each participant, and compared across participants. The hypothesis that more information leads to better performance would predict that no matter the overall speed of a participant, they should perform relatively faster with more available information. We found no significant differences between information volume conditions ($F(2,46) = 0.49$, $p = 0.62$, $\eta^2 = 0.02$). Neither did the data trend in this direction; conversely, the fastest condition was actually the low information volume condition (see Figure 5).

As a comparison, the Ideal Observer Model was designed to perfectly integrate all intelligence updates. The purpose of this model was to discover if additional information objectively conferred an advantage in decision-making. For example, it was possible that the timing of the task and the nature of the updates dictated that only using the first update to assign a unit, and then searching the area after arriving was as good as or a better strategy than processing several updates to make a better guess at the target location. If the Ideal Observer Model did not perform increasingly better with more information (5 and 9 update conditions) than less information (1 update condition), this would indicate that the human participants were not necessarily overwhelmed by and unable to use the additional information; but that they were making a possibly rational decision to only use a single update, no matter how many were available. However, if the ideal observer did perform better with more information, this would indicate that the human participants were unable to take advantage of the benefits of additional information.

The ideal observer model was run on the exact set of intelligence updates received by each participant. As with the human data, the rates were converted to z-scores and then averaged (see Figure 5). Contrary to the human data, the ideal observer performed much faster with additional intelligence updates. A two way mixed ANOVA confirms this result; the interaction between the information volume (low, medium, or high) and the type of data (human vs. ideal observer) is significant ($F(2,90) = 19.44$, $p < 0.0001$, $\eta_p^2 = 0.30$).

The ideal observer model appears to have diminishing returns as it asymptotically approaches a performance ceiling. With sufficient relevant information, even imperfect information, the ideal observer gets closer to being fully deterministic; much like Laplace's Demon where perfect information enables perfect predictability – with the only constraint being the time to move the units.

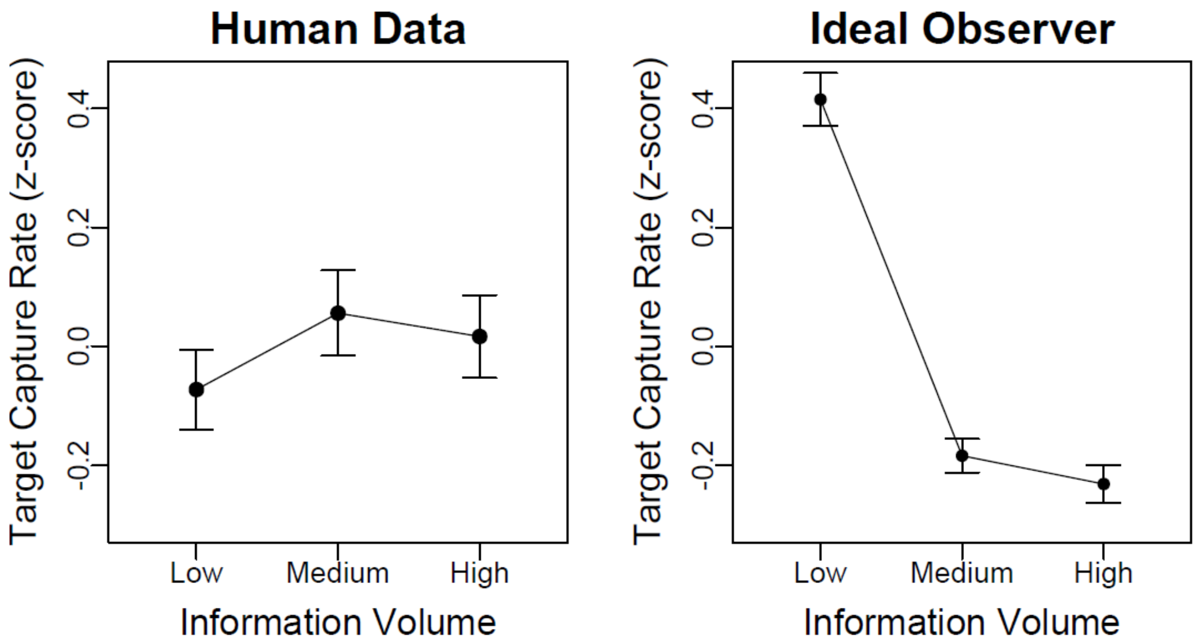

Figure 5. Comparison of experimental results from human participants and simulated results from the ideal observer model. Larger z-scores represent relatively slower target capture times, while smaller z-scores represent relatively faster target capture times. Humans performed better in the low information condition whereas performance of a data fusion algorithm (Ideal Observer Model) improves with increasing amounts of information. Error bars represent standard error of the mean.

Discussion

In our study we manipulated the volume of task-relevant information provided to participants and measured the resulting time to find and capture high value targets. We anticipated one of two outcomes: 1) More information leads to better performance ("More is More"), or 2) More information leads to worse performance ("More is Less"). We found, however, that increasing the volume of task-relevant information did not impact human performance on the task. This outcome was not one of those hypothesized, and might be thought of as "More is the Same." In contrast, an ideal observer, which perfectly integrated all information provided, performed much faster with increasing information ("More is More"). The results from the Ideal Observer Model demonstrate that in this task, computational performance can be improved by integrating all available information. Human performance data showed neither improvement nor detriment with increasing information volume; this indicates that for the current task human participants may have been at their limits in integrating or fusing information.

Our results provide evidence of a limited human information fusing capability – individuals are not always able to take advantage of all information provided to them, even when this information is useful to the task at hand. These findings indicate that in C2 environments, caution should be exercised when attempting to make all information available to all personnel. In addition, these results make a strong case for continued research into effective decision-support tools that can assist in information synthesis and disambiguation.

We find evidence that in some high-information volume environments, human decision-making performance can be surpassed by an information fusion algorithm (such as our Ideal Observer Model). However, when automated decision-support tools are incorporated into the military C2 context, human supervisory control is still clearly a requirement. Future work might explore the optimum interaction between automated fusion algorithms and human cognitive fusion in similar simulated experimental C2 tasks.

Conclusion

The purpose of this work was to begin to evaluate the impact of the transition to NEO, made possible by interconnected systems and communications, upon the human decision maker. In particular, we assessed how the rapidly increasing amounts of available information help or hinder decision-making performance. We systematically examined this research question by developing a simple military-relevant task focused on HVT capture; this task allowed for direct manipulation of information volume and straightforward measurement of human performance. In addition, the simplified task also allowed us to develop an Ideal Observer Model for optimal task performance, which we compared to actual human performance. We found that, while more information allowed an optimal observer to perform faster, this was not the case for human participants. We find evidence that more information is not necessarily better for human decision-making, consistent with the literature on information overload. There are clear limitations to our study; for example, our participants completed a single task with simplifications from a real C2 environment. However, this experiment is the first step in a program of research focused on studying the effects of networked operations upon the human decision maker. Our next phase of study includes expanding this task to allow two roles, an Intelligence Officer and an Operations Officer, to perform different functions while interacting to successfully complete the mission. In addition to increasing the complexity of the task, this will also introduce issues of team dynamics, communication, and trust, allowing for richer data sets to investigate the wide variety of questions surrounding human decision-making performance in networked operational environments.

References

Alberts, D., & Garstka, J. (2004). *Network Centric Operations Conceptual Framework Version 2.0*; Evidence Based Research, Inc.: Vienna, VA.

Blasch, E., Bosse, E., & Lambert, D. A. (Eds.). (2012). *High-level Information Fusion: Management and Systems Design*. Artech House.

Chewning, E. G., & Harrell, A. M. (1990). The effect of information load on decision makers' cue utilization levels and decision quality in a financial distress decision task. *Accounting, Organizations, and Society*, *15*, 527-542.

Committee on Network Science for Future Army Applications (CNSFAA) (2005). National Research Council, Network Science. Washington, DC: The National Academies Press.

Edmunds, A., & Morris, A. (2000). The problem of information overload in business organisations: A review of the literature. *International Journal of Information Management*, *20*, 17–28.

Geisler, W. S. (2006). Ideal observer analysis. In L. Chalupa & J. Werne (Eds.), *The Visual Neuroscience* (pp. 825-837). Boston: MIT Press.

Nadav-Greenberg, L., & Joslyn, S. L. (2009). Uncertainty forecasts improve decision making among nonexperts. *Journal of Cognitive Engineering and Decision Making*, *3*, 209-227.

Swan, J. M., & Hennig, J. (2012). From data to decisions. *Army AL&T*, 8-12.