# Effects of Information Availability on Military Decision-Making: Performance, Trust, and Situation Awareness

Laura R. Marusich

U.S. Army Research Laboratory

Emrah Onal

SA Technologies

Michael S. Yu

Carnegie Mellon University

Jonathan Z. Bakdash

U.S. Army Research Laboratory

James Schaffer, John O'Donovan, Tobias Höllerer

University of California, Santa Barbara

Norbou Buchler

U.S. Army Research Laboratory

Cleotilde Gonzalez

Carnegie Mellon University

Author Note

Correspondence concerning this article should be addressed to Laura Marusich, U.S. Army Research Laboratory, Bldg 91012 Station Ave, Fort Hood, TX 76544, laura.r.marusich.ctr@mail.mil

# Effects of Information Availability on Military Decision-Making: Performance, Trust, and Situation Awareness

Technological advances drive information growth and transform the ways that people share information (Gleick, 2011). To leverage information growth and sharing, with the goal of improving their military operational capabilities, the United States and other NATO countries continue to invest heavily in Network Enabled Operations (NEO). NEO involves highly connected information and communication systems, allowing fast and flexible sharing of information within and between units and across geographical locations. With such a dramatic shift to extensive networking, there is a real need to understand the conduct of NEO at the human cognitive level, and in particular to understand the impact that operating in complex, information-rich environments has upon human decision making performance. In time-stressed situations common to the Command and Control (C2) environment, the performance of the entire networked organization can be constrained by the ability of any single member of the staff to process information in a timely manner.

The transformation of the U.S. and NATO countries to NEO has proceeded under a conceptual framework 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 leads to better situation awareness (SA), decision-making, and ultimately, mission outcomes. In essence, the increase in information available to commanders and their staff is expected to result in better decision-making based on enhanced SA (CNSFAA, 2005). However, the tenets of this framework are in need of systematic study, especially in relation to decision-making and human information processing in C2 environments. For instance, there may be situations in NEO where increased information sharing increases 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 or redundant information (the noise). Even when information sharing results in increased availability of relevant information, the sheer volume and rapid pace of information received and readily accessible through networked systems can be overwhelming. In other words, effects of task workload and time constraints in the environment on human decision making situations such as C2, will directly depend on the individuals' cognitive processing abilities (Gonzalez, 2004, 2005b)

Evidence from research on dynamic decision-making tasks suggests that, contrary to the tenets of NEO, more information does not necessarily lead to better decisions (e.g. Gonzalez, 2005a); and many times a closer and more robust networked force may result in poorer shared SA (Saner, Bolstad, Gonzalez, & Cuevas, 2009). For example, Nadav-Greenberg and Joslyn (2009) asked their 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 given the expected overnight temperature. The researchers manipulated what additional information items the participants received: no additional information, the lower bound of the 80% confidence interval (CI) 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. In a different study, participants predicted a firm's financial distress on the basis of 4, 6, or 8 different information cues (Chewning & Harrell, 1990). The researchers found that approximately one third of their participants showed signs of information overload, choosing not to use all the available information cues and making less consistent decisions overall.

Similar effects have been found in simple judgment tasks: participants given additional relevant information about cities in the United States and Germany were not more accurate in ranking the cities by size (Goldstein & Gigerenzer, 2002). In addition, participants were less accurate in predicting the outcome of professional basketball games when they received additional information about the team name than when they lacked that information (Hall, Ariss, & Todorov, 2007).

The results from past studies suggest that more information may not always help and may in fact result in cognitive overload. However, most decision-making tasks used in past studies are less typical of the demands of the C2 environments (see Gonzalez, 2004; 2005b for an exception); in addition many of these tasks have lacked an objective ground truth with which to compare participants' decisions or have included conditions where additional information provided is redundant or irrelevant to the task. Understanding of the consequences of operating in an information-rich, time-stressed military operations environment is a critical challenge for human performance, and a high priority for the Office of the Secretary of Defense's 'Data to Decisions' initiative. This initiative calls for ways to manage the complexity of the information environment in ways that "enable faster, better decisions while reducing information overload" (Swan and Henning, 2012). The goal of our first experiment is to examine how the volume of task-relevant information influences objectively measured performance in a military relevant decision-making task. In Experiment 1 we examine the impact of task-relevant information volume upon the decision-making performance of single participants operating individually in a C2 environment.

A possible way to reduce possible detrimental effects of information overload is to distribute information processing tasks across the network – allowing separate people to process and act upon different sets of information (see, for example, Kozlowski, Gully, Nason, & Smith, 1999; Salas, Goodwin, & Burke, 2009) . Such distribution also increases dependencies, requiring each person in the network to rely on the performance of others.

This willingness to accept vulnerability to others, for a potential common benefit, is defined as trust (Mayer, Davis, & Schoorman, 1995). Increased transparency and awareness of information across individuals within a network may serve to increase trust, as one can verify another person's work (e.g., Saner et al., 2009; Martin, Gonzalez, Juvina and Lebiere, in press). On the other hand, such verification may itself be a sign of distrust, and observed mistakes in another person's work may be over-weighted. Similarly, if judgments made by others in the network are flawed, it may be unclear where the source of these errors may arise, making it difficult to distinguish between good judgments made from poor information and poor judgments made from good information; this makes assessment of interpersonal trust difficult. Experiment

2 builds upon Experiment 1, using two-player teams instead of single players. The goal of our second experiment is to explore the effects of information volume and information reliability upon decision-making performance in distributed operations, and additionally upon interpersonal trust between teammates in such environments.

**Simulated C2 Task**

Our C2 task used in these experiments is focused on a single goal of finding and capturing high value targets (HVTs) within a given area of operations (AO). Participants received text-based intelligence information about the location of HVTs, and they had control over the movements of four identical assets that could be assigned to travel to any location on the map to capture HVTs. Friendly unit icons based on standard military symbology (DoD, 2008) were used to represent the assets. An HVT was considered "captured" if a unit icon was moved to the same map location as the HVT. Objective decision-making performance was operationalized as the time to capture the hidden targets. Experiment 1 was a single-player C2 task: Participants controlled unit movements to capture targets using intelligence reports with varying amounts of information, all of which was relevant. Experiment 2 was a two-player C2 task: Participants worked collaboratively in pairs, one responsible for unit movements and the other for intelligence reports. In Experiment 2, we examined decision-making given source reliability information and limited versus self-selected sharing of information. In addition, we investigated the relationship between decision-making and self-reported measures of informational and interpersonal trust and SA. The two versions of the task are described in detail below.

**Single-player task.** In the single-player task used in Experiment 1, the display contained a map of the AO, a text box which displayed incoming intelligence information about the location of HVTs, a text box which showed spot reports from the four units, a progress bar indicating how many targets had been captured, and checkboxes for each target which participants could use to mark which targets had been captured (see Figure 1). The map was divided into a two-dimensional grid. The x-axis of the map grid was marked with numbers 1 through 14, and the y-axis was marked with letters A through N. This common framework was used for referencing to particular locations on the map.

During the task, new HVTs activated somewhere in the AO every 15 seconds and remained in the same location until captured. When a new HVT activated, participants received intelligence information about the possible location of that target in the “Intel” text box (e.g. “HVT 1 sighted at D2”), but they were not shown the HVT on the map. Using this information, the participant could click on any of the four unit icons and assign them to travel to map 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). As a result of these constraints, 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).

**Two-player task.** In the two-player task used in Experiment 2, each of the two players was assigned to either the role of the Intelligence Officer or the Operations Officer. The display elements visible to participants depended on their assigned role and experimental condition. The goal of the task was the same as in the single-player version: to use text-based intelligence information to find and capture HVTs.

*Intelligence Officer Role (Intel).* The responsibilities of the participant assigned to the Intel role were to read and process the incoming intelligence updates and to use this information to mark probable HVT locations on the grid map. This was done by dragging target icons to the map (see Figure 4), and these icons were automatically shared with the other player in all conditions.

*Operations Officer Role (OPS).* The responsibilities of the participant assigned to the OPS role were to use the information on probable target locations based on available information to assign the four units to find and capture the targets and to keep track of which targets had been captured. As in the single-player version of the task, this was done by clicking on individual units and either clicking on the grid space to assign a goal location or typing in the grid location to assign it. The yellow and red arrows indicating unit movement were visible to both players in all conditions.

In this version of the task, the intelligence updates were provided by two sources: Source A and Source B. In each round of game play, one of these sources was randomly assigned to be 90% likely to be accurate, and the other would only be 10% accurate. As in the single-player version, if the information was not accurate, it was only off by one square, either horizontally or vertically. Four updates appeared for each target, with two from Source A and two from Source B. The order of presentation reversed on each target (AABB, BBAA, AABB,etc.). Updates for the same target were presented every 4 seconds, and new targets activated every 18 seconds. Another additional display element in the two-player version of the task was a chat window, which allowed the two participants to communicate directly with one another if they chose to do so.

Both the single- and two-player versions of this experimental task entail clear simplifications from a real world C2 scenario. For example, we limit the interaction to one or two roles, rather than the multiple interacting roles of a command staff. In addition, this task was focused on one well-defined mission, to find and capture HVTs, rather than a more complex combination of operations and events. However, the experimental design conferred the benefits of controlled manipulation of relevant factors, allowing for systematic study of our variables of interest. For example, in Experiment 1 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 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. 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, target capture times served as a useful quantitative measure of task performance.

**Hypotheses**

In Experiment 1, we used the single-player version of the task, and manipulated the number of location updates presented to the participant about each HVT. Based on the literature, we expected either of two outcomes:

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.

In Experiment 2, we used the two-player version of the C2 task. We investigated the effects of information availability by manipulating whether players had access to the information primarily used by their partners. As in Experiment 1, we expected possible divergent outcomes. It is possible that making all information available to both partners would allow for improved performance over a more limited-information condition, as predicted by the tenets of NEO.

Conversely, this access to additional information might lead to information overload and poorer task performance. Similarly, we anticipated one of two possible outcomes of the effect of information availability upon interpersonal trust. The transparency of shared information may promote trust; however the corresponding opportunity to second guess a partner's decisions may lead to decreased trust. In addition to information availability, we manipulated the knowledge provided about the reliability of the information provided – either providing a correct reliability report, an incorrect reliability report, or no reliability report. We hypothesized that task performance would be best with correct information, followed by no information, and worst with incorrect information.

## Experiment 1 – HVT Single Player

Experiment 1 was designed to assess the effects of different amounts of information upon task performance in a military-relevant C2 task. The tenets of NEO would predict that more information should lead to better performance, represented in this task by faster target capture times. In addition, this experiment provided insight into individual performance on this task, informing the design and analysis of Experiment 2, in which two players completed the task in pairs.

### Method

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

**Experimental Design.** In this experiment, the independent variable was the number of intelligence reports (1, 5, or 9) presented to the participant about the location of a single HVT.

All updates about a single target appeared within a 16-second window, with new targets activating every 15 seconds. The number of updates varied by experimental block, and the order of blocks was counterbalanced across participants. The dependent variable was the time it took participants to capture an HVT after presentation of the first intelligence report about that HVT.

**Procedure.** The study took place in a dimly lit experimental room on a single monitor, Windows-based 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 using both text and diagrams. 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 of 18 HVTs each.

**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 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.

The output of this model was used during data analysis only; it was not shared with the participant. This model is useful in that it defines ideal information fusion performance 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 the data our human participants actually generate.

**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, these rates were converted to standardized z-scores for each participant. We then calculated the average z-score rate in each information volume condition for each participant, and compared across participants. The tenets of NEO would suggest that no matter the overall speed of a participant, they should perform relatively faster with more available information. In contrast, we found no significant differences in capture time between the various 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, an ideal observer model was designed, subject to the same constraints of the task, that was able to perfectly integrate all intelligence updates. The purpose of this model was to discover if additional information objectively conferred an advantage. The ideal observer model was run on the exact set of intelligence updates received by each participant. If the ideal observer model did not perform better in the 5 and 9 update conditions than the 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 in the 5 and 9 update conditions, this would indicate that the human participants were unable to take advantage of the benefits of additional information.

As with the human data, the rates were converted to z-scores and then averaged (see Figure 5). In contrast to the human data, the ideal observer performed much faster with additional intelligence updates, relative to the lowest information condition. There is some evidence of diminishing returns in model performance as the number of updates increases from 5 to 9, indicating that in this task an intermediate number of updates is often sufficient to determine target location. More updates allow even more certainty of target location, but also increase the probability of redundant information.

**Discussion**

In this 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.

This result appears to challenge the tenets of NEO, which holds that increased information should improve the efforts of mission command. Indeed, research in psychology suggests that people are limited by how much information they can process, given limited working memories (Anderson, Reder & Lebiere, 1996; Baddeley, 1992; Miller, 1956). These limitations, in turn, have been linked to our ability to perceive, comprehend, and make projections about the environment we are facing – referred to more generally as SA (Endsley, 1995)

However, the promise of an integrated network is not limited to information availability, but also to better, more distributed information processing. In particular, multiple people integrated into the efforts of Mission Command allows for both increased capacity and flexibility in evaluating the environment. Yet, while dividing information processing tasks can reduce the cognitive demands on individuals in the network, it may also introduce the risk that the right information might not reach the right people at the right time, particularly when information reliability varies. This raises the question of how to most effectively share this information across the network. Additionally, interpersonal dynamics become increasingly important. Whether distributed efforts can be consolidated to coherent and effective solutions may depend upon the interpersonal trust between operators. When trust is low, judgments made by colleagues might be discounted or ignored, leading to either missing, redundant, or delayed action.

## Experiment 2 - HVT Two-Player

In this study, we explored issues of interpersonal dynamics and information availability using the two-player version of the C2 task. We manipulated both information availability and source reliability information and assessed the impact on human performance, SA, and interpersonal trust.

SA forms the basis for decision-making and task performance (Endsley & Garland, 2000; Strater & Bolstad, 2008; Strater, Endsley, Pleban & Matthews, 2001). While it is possible for an individual with an incomplete or inaccurate understanding of the situation to make the right decision and perform well by chance, a good outcome is far more likely if that individual has an accurate picture of the situation. In this study, we measure SA to have a better understanding of the behavior exhibited by the individual, including trust and performance. SA measurement also helps interpret findings by providing insight into the individual's cognitive state preceding decision-making. As suggested by Study 1, one potential risk of increasing available information is that we might actually decrease rather than increase SA if there is information overload. In this study, we therefore investigate the effects of information transparency and information quality on judgments of trust in the C2 context.

## Method

**Participants.** Twenty-eight complete pairs of participants (56 participants total – 24 male, 32 female) volunteered for this study through the subject pool in the Psychology Department of the University of California, Santa Barbara. Participant ages ranged from 18 to 31 years.

**Experimental Design.** There were two manipulations in this study. The first was a between-subjects manipulation, whether the role-specific information was shared or limited to the relevant role (see Figure 6). In the limited condition, only the Intel player was able to see the Intelligence updates and source reliability information (region a1 of Figure 6), and only the OPS player was able to see the spot reports about unit movement and target captures (region a2 of Figure 6). In the shared condition, all of these elements were visible to both players.

The second manipulation was within subjects, and it determined what information was provided to the participants about source reliability. In the "none" condition, no information was provided. In the "congruent" and "incongruent" conditions, an additional display element (region b of Figure 6) was present which provided information about source reliability. In the "congruent" condition, this information correctly indicated which source was 90% accurate and which was 10%. In the "incongruent" condition, the information was reversed. For example, if Source A was actually 90% accurate in that round, the display would show Source B as 90% accurate and Source A as 10% accurate.

In all conditions, we collected data on task performance (measured by target capture times), subjective SA, and players' trust in their partners (both measured by questionnaires completed after each round of the task, described below).

**Questionnaires.**

*Pre-test questionnaire.* Participants completed this questionnaire before beginning the experimental task. The goal of this questionnaire was to gather general demographic information, gauge participants' propensity to trust, and assess their familiarity with the concepts presented in the main study. Most of the questions used a three point Likert scale for responses.

*Round questionnaire.* This questionnaire was given to participants after each experimental block (see Appendix). The first part of the questionnaire was the Mission Awareness Rating Scale (MARS), a subjective measure designed to assess SA content and SA workload adapted from the Crew Assessment Rating Scale (CARS; McGuiness & Foy, 2000) to the military domain (see Appendix). It has been validated in a series of experiments by Matthews, Beal, & Pleban (2002). MARS assesses participants' perceptions of their SA. It was selected due to its origins in the military domain and its unobtrusiveness for this experiment.

MARS consists of 2 subscales. Each subscale consists of four questions, on a 6-point scale, that address the three levels of SA as defined by Endsley (1988) – perception (identify), comprehension (understand), and projection (predict). A fourth question is about mission goals. The first subscale is used to assess SA content, while the second subscale is used to assess SA workload.

MARS questions were adapted to our experiment by making minor wording changes to the questions. In addition, three more questions about the teammate (awareness of the teammate's activities, sharing decision-making, and communicating) were added to assess shared SA.

The second part of the round questionnaire assessed the participant's self-reported trust in the other player. This relatively simple assessment of trust was chosen to avoid overwhelming participants, as this questionnaire was presented on three occasions.

*Post-test questionnaire.* After the main study, participants completed a short post study questionnaire. The purpose of this questionnaire was to gauge participants' understanding of the game, their perspective on the utility of the UI components, and to assess each player's feeling about their partner player's competence.

**Procedure.** Each pair of participants was randomly assigned to either the Shared or Limited information, and roles were randomly assigned within pairs at the start of each session. The two members of a pair completed the study on single monitor computers in separate experimental rooms. Participants first completed the pre-study questionnaire. They then performed a self-paced tutorial that provided an overview of the purpose of the task, the responsibilities of each role, and that allowed them to step through each of the actions that would be required of them in the task. Participants then completed three test blocks corresponding to the three source reliability conditions (Congruent, Incongruent, None). The order of these blocks was counterbalanced across pairs. Each block consisted of a practice run of four targets, followed by a test run of 18 targets. After each block, participants completed the round questionnaire, and after all three blocks they completed the post-study questionnaire.

## Results

**Access to Information (limited/shared) Manipulation.** A two-way mixed ANOVA showed no significant main effect of the between-subjects Limited/Shared manipulation upon task performance, measured by target capture times.

There was a main effect of the Limited/Shared manipulation upon SA levels, measured by the MARS questionnaire ($F(1,54) = 8.29$, $p = 0.006$; see Figure 7), with lower SA scores in the Shared condition ($M = 3.04$, $SD = 0.40$) than in the Limited condition ($M = 3.32$, $SD = 0.33$).

In addition, there was a main effect of the Limited/Shared manipulation upon self-reported trust in one's partner ($F(1,54) = 6.40$, $p = 0.014$; see Figure 8), with lower trust scores in the Shared condition ($M = 3.42$, $SD = 0.66$), than in the Limited condition ($M = 3.76$, $SD = 0.48$).

**Source Reliability Information Manipulation.** A two-way mixed ANOVA showed a significant main effect of source reliability information upon task performance, measured by target capture times ($F(2,52) = 7.96$, $p = 0.001$). Post-hoc Tukey HSD indicated significantly slower target capture times in the Incongruent condition ($M = 16.97$, $SD = 6.93$) than in either the None ($M = 13.24$, $SD = 5.19$) or the Congruent ($M = 11.20$, $SD = 5.63$) conditions (see Figure 9).

In addition, there was a main effect of source reliability information upon SA, as measured by the MARS questionnaire ($F(2,108) = 7.14$, $p = 0.001$); see Figure 7). Post-hoc Tukey HSD showed that SA scores were significantly lower in the incongruent condition ($M = 3.04$, $SD = 0.45$) than the congruent condition ($M = 3.33$, $SD = 0.54$).

The main effect of source reliability information upon self-reported trust in one's partner was marginal ($F(2,108) = 2.49$, $p = 0.088$), with the lowest levels of trust reported in the Incongruent condition (see Figure 8).

There were no significant interaction effects between the Shared/Limited manipulation and the source reliability information manipulation upon any of the above dependent variables.

**Discussion**

We found that both the between-subjects Limited/Shared and the within-subjects source reliability information manipulation affected the participants' experience on the simulated C2 task. While participants in the Shared condition did not exhibit significantly slower target capture times than those in the Limited condition, they did report lower levels of SA, as well as lower levels of trust in their partners. It seems to be the case that when users have the opportunity to see the information their partners are using to perform their duties, the users are able to "second-guess" their partners' decisions and as a result, feel less trust in their partners. This additional information may also cause a degree of overload in users, especially if they spend substantial time using it to check their partners' work, resulting in lower levels of SA.

As discussed above, participants' SA was measured using MARS, a subjective measure of SA based on self-ratings. While less intrusive to participants and the experiment, self-ratings have a number of aspects that may impact how we interpret our findings. MARS was administered as a post-trial questionnaire reflecting participant SA at the end of each round of the task, and people are generally poor at reporting detailed information about past mental events (Endsley, 1995). Also, people may not be very good at rating their own SA accurately. In our study, reported SA was found to be significantly lower under the Shared condition. It is possible that actual participant SA was not lower, but rather participants' confidence in their own SA was lower. This might have been driven by the lower trust in partner or having to process additional information, lowering their confidence and giving a perception of lower SA. This is partly supported by the fact that target capture times were not significantly slower in the Shared condition. For future studies we will consider integrating objective measures of SA into the experiment, such as Situation Awareness Global Assessment Technique (SAGAT; Endsley 1988), for a better understanding of the observed behavior.

One consequence of the design of this task is that in the Shared conditions, the OPS players can go beyond simply attending to the additional information available in the Shared condition to actually using the information to perform both roles, ignoring the work of their (Intel) partner partially or completely. In cases where the Intel player is performing poorly, this may be an effective strategy. However, it can easily result in worse overall performance, as a result of information overload and task overload. We examined the data for evidence of this type of behavior in the OPS players. We noted the time at which the Intel player communicated a target location to the OPS player, through either the map interface or the chat window. We then noted the time that the OPS player first assigned a unit to capture that target. If the OPS action preceded the Intel action, this provided evidence that the OPS player was appropriating the Intel player's duties. We found, relative to the baseline timing difference in the Limited condition, that timing differences in the Shared condition were not significantly less (see Figure 10). However, there was a wider range of values, with large variability even within a single round of play of some pairs. This indicates that in some pairs, the OPS player both acted much earlier than the Intel player and much later, as compared with pairs in the Limited condition. In fact, we find that the magnitude of this variability within a single round of game play predicts task performance (target capture times). That is, in rounds with high variability in the time difference between actions, participants took longer to capture targets (see Figure 11). The implication is that the strategy of taking over for one's partner in this task may result in faster performance for some trials, but the additional task demands will cause overload and lead to much slower performance on many other trials.

This deeper analysis shows that there is no one condition that performs the best in all cases. As follow on research, this observation may point to an opportunity for exploring adaptive systems that change the UI based on underlying conditions. An adaptive system may be able to exploit the fact that in some cases Shared condition can offer a benefit (taking over for one's partner for faster performance for certain scenarios) whereas in general the Limited condition has better overall performance. An interesting research question would be to investigate whether the benefits of an adaptive interface outweighs the additional complexity and overhead introduced by it.

The team nature of this task allowed us to explore effects upon interpersonal trust, i.e. trust in one's partner. We were also interested in manipulating reliance upon and trust in information sources and assessing the effects upon task performance. We found that providing participants with false information about the reliability of the two information sources (the Incongruent Condition) resulted in slower target capture times, lower self-reported SA, and possibly lower self-reported trust in players' partners. We did not find evidence that having correct information about the source reliabilities (the Congruent Condition) resulted in better performance, higher SA or higher trust, as compared to having no information about source reliabilities. This indicates that, at least in this task, the cost of having false information about the accuracy of information sources is larger than the benefit of having true information.

## Conclusions

As technological advances and the military's transition to Network Enabled Operations give rise to an explosion in the amount of information available to individual Soldiers, it is critical to understand the impact of information volume upon human decision-making, SA, and interpersonal trust. To this end, we designed two empirical studies that explored information availability in a military relevant decision-making task. The first of these studies investigated how varying the amount of task-relevant information presented to single participants influenced the time it took them to find and capture targets. The second study added the element of team dynamics; two interacting participants worked together to complete a version of the same high value target task.

The general findings from these two studies indicate that increasing the amount of available task-relevant information is not necessarily beneficial to human decision-making performance, in agreement with previous findings (e.g. Chewning & Harrell, 1990; Goldstein & Gigerenzer, 2002; Gonzalez, 2005; Hall, Ariss, & Todorov, 2007; Nadav-Greenberg & Joslyn, 2009; Saner et al., 2009). Furthermore, Experiment 2 provides evidence that access to increased information may reduce reported SA and trust in team members. These are important findings, particularly in the military domain, where decisions are made in time-stressed, life-or-death situations. In these environments, human decision makers may be easily overwhelmed with large volumes of information, even when that information is potentially useful. Consequently, there is ample opportunity for the development of decision-support systems that can assist in fusing, integrating, and presenting useful information to the human user. 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.

**References**

Alberts, D., & Garstka, J. (2004). Network centric operations conceptual framework version 2.0. *US Office of Force Transformation and Office of the Assistant Secretary of Defense for Networks and Information Integration, US Department of Defense, Tech. Rep*.

Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology*, *30*, 221–256.

Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559.

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.

Colquitt, J. A., Scott, B. A., & LePine, J. A. (2007). Trust, trustworthiness, and trust propensity: a meta-analytic test of their unique relationships with risk taking and job performance. *Journal of Applied Psychology*, *92*, 909.

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

DoD (2008). MIL-STD-2525C: Common Warfighting Symbology: U.S. Department of Defense Interface Standard.

Dirks, K. T. (1999). The effects of interpersonal trust on work group performance. *Journal of Applied Psychology*, *84*, 445.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors and Ergonomics Society 32nd Annual Meeting* (pp. 97–101). Santa Monica, CA: Human Factors and Ergonomics Society.

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

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

Endsley, M. R., & Garland, D. J. (2000). *Situation awareness analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum.

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

Gleick, J. (2011). *The Information: A History, a Theory, a Flood*. Random House LLC.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*, 75–90.

Gonzalez, C. (2004). Learning to make decisions in dynamic environments: Effects of time constraints and cognitive abilities. *Human Factors, 46*, 449–460.

Gonzalez, C. (2005a). Decision support for real-time dynamic decision making tasks. *Organizational Behavior and Human Decision Processes*, *96*, 142–154.

Gonzalez, C. (2005b). The relationship between task workload and cognitive abilities in dynamic decision making. *Human Factors, 47*(1), 92-101.

Hall, C. C., Ariss, L., & Todorov, A. (2007). The illusion of knowledge: When more information reduces accuracy and increases confidence. *Organizational Behavior and Human Decision Processes, 103*, 277-290.

Kozlowski, S. W. J., Gully, S. M., Nason, E. R., & Smith, E. M. (1999). Developing adaptive teams: A theory of compilation and performance across levels and time. In D. R. Ilgen &

E. D. Pulakos (Eds.), *The changing nature of work performance: Implications for staffing, personnel actions, and development* (pp. 240-292). San Francisco: Jossey-Bass.

Levin, D. Z., & Cross, R. (2004). The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. *Management Science*, *50*, 1477–1490.

Martin, J. M., Gonzalez, C., Juvina, I., & Lebiere, C. (in press). A description-experience gap in social interactions: Information about interdependence and its effects on cooperation. *Journal of Behavioral Decision Making.*

Matthews, M. D., Beal, S. A., & Pleban, R. J. (2002). *Situation awareness in a virtual environment: Description of a subjective assessment scale*. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review*, *20*, 709–734.

McAllister, D. J. (1995). Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, *38*, 24–59.

McGuinness, B., & Foy, L. (2000). A subjective measure of SA: the Crew Awareness Rating Scale (CARS). In D. B. Kaber and M. R. Endsley (Eds). *Proceedings of the First Human Performance, Situation Awareness, and Automation Conference.* Savannah, Georgia.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

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

Salas, E., Goodwin, G. F., & Burke, C. S. (Eds.). (2009). *Team effectiveness in complex organizations. cross-disciplinary perspectives and approaches.* New York, NY: Taylor & Francis.

Saner, L. D., Bolstad, C. A., Gonzalez, C. & Cuevas, H. M. (2009). Measuring and predicting shared situation awareness in teams. *Journal of Cognitive Engineering and Decision Making, 3*, 280-308.

Strater L. D., & Bolstad C. A. (2008). Simulation-based Situation Awareness Training. In D. A. Vincenzi, J. A. Wise, M. Mouloua, & P. A. Hancock (Eds.), *Human Factors in Simulation and Training*. (pp. 129-148). New York: Lawrence Erlbaum Associates.

Strater, L. D., Endsley, M. R., Pleban, R. J., & Matthews, M. D. (2001). *Measures of platoon leader situation awareness in virtual decision-making exercises*. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Zaheer, A., McEvily, B., & Perrone, V. (1998). Does trust matter? Exploring the effects of interorganizational and interpersonal trust on performance. *Organization Science*, *9*, 141–159.

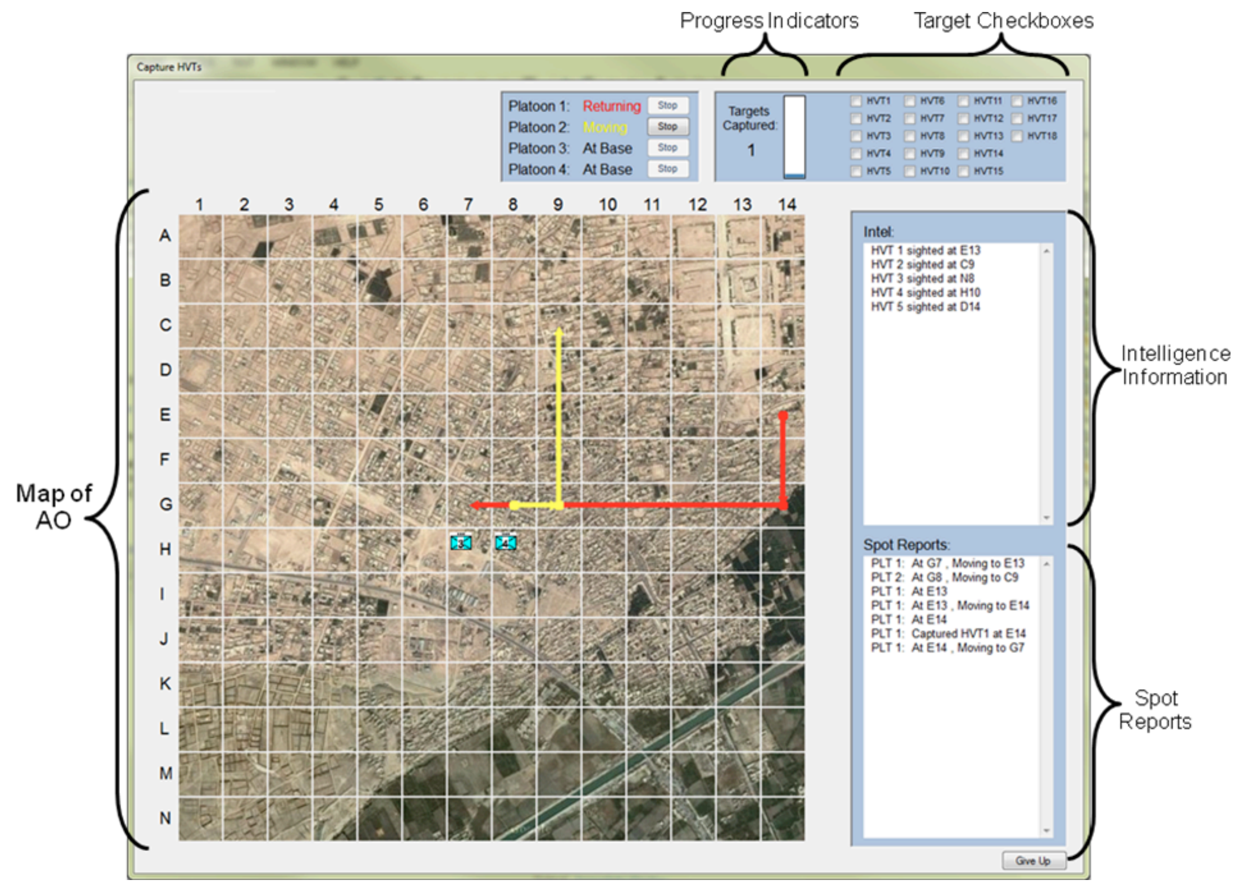


*Figure 1*. Screen capture of the single-player task display used in Experiment 1.

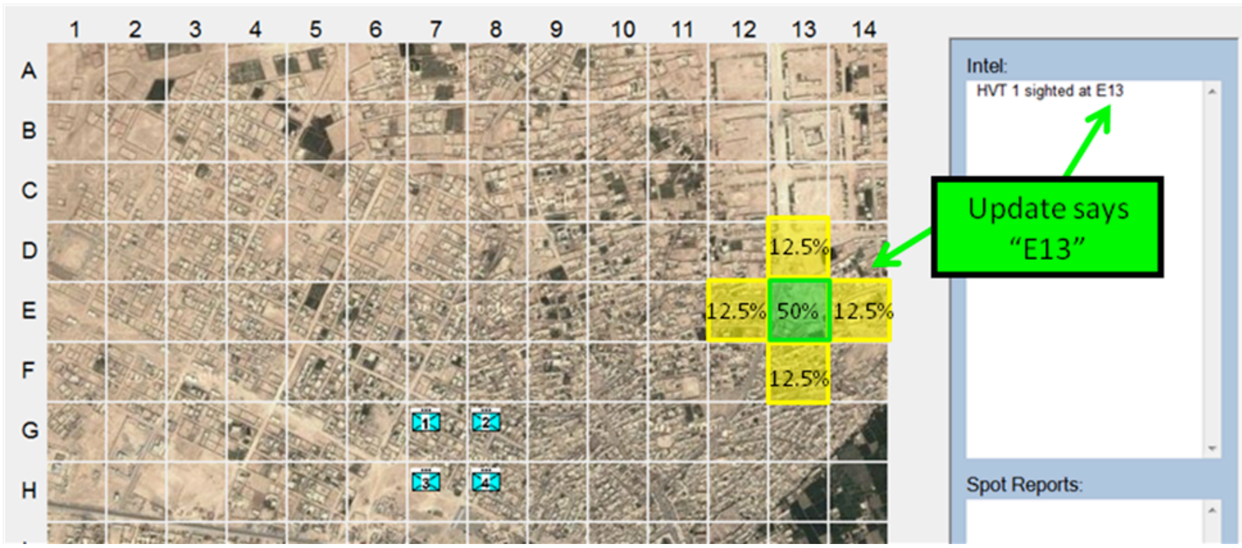


*Figure 2.* Illustration of possible HVT locations and their respective probabilities, given a single intelligence update. This illustration was shown to participants during the tutorial phase, but it was not part of the experimental display.

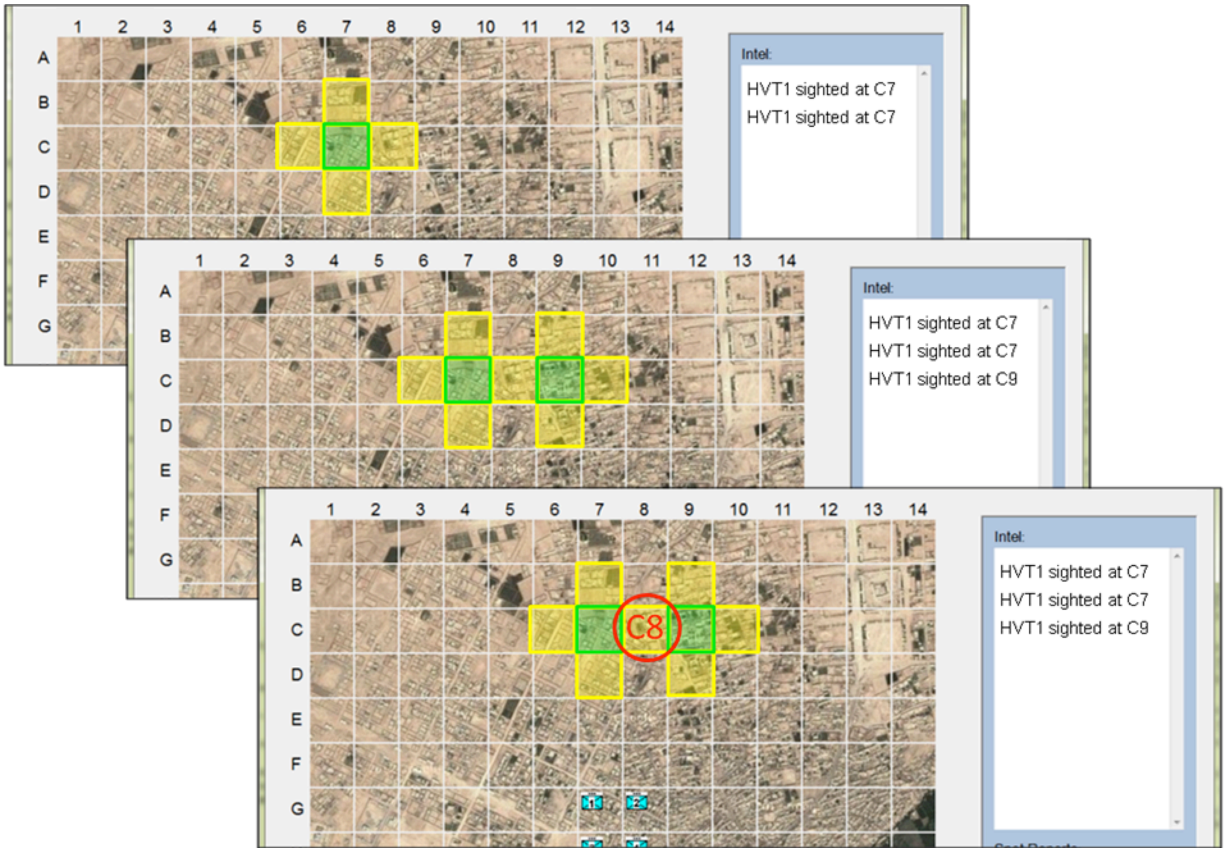


*Figure 3.* Illustration of integrating multiple intelligence updates to determine the true target location. Given the possible target locations associated with each of the two unique updates, the only possibility is C8. These illustrations were not part of the experimental display.

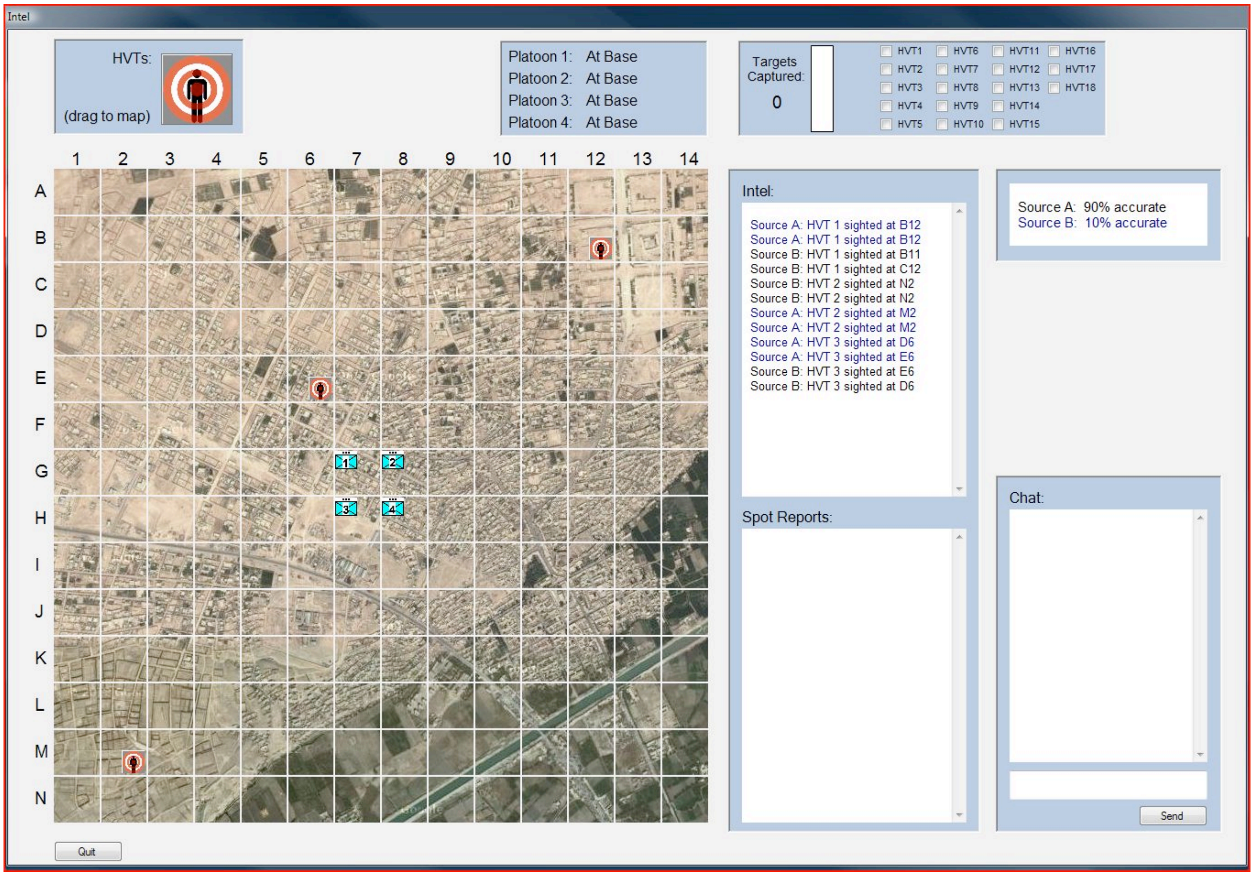


*Figure 4*. Screen capture of the Intel player’s display in the two-player task used in Experiment 2.

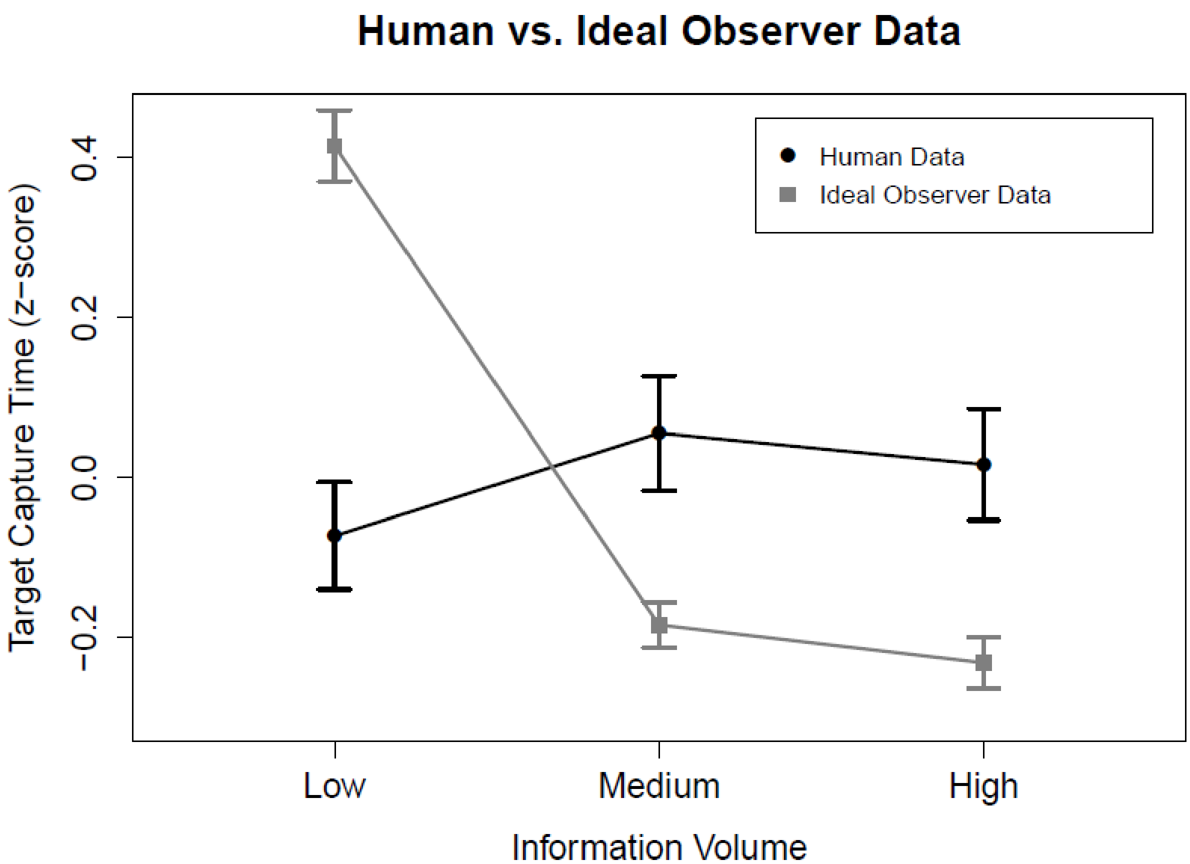


*Figure 5*. Comparison of results from human participants and simulated results from the ideal observer model in Experiment 1. Larger z-scores represent relatively slower target capture times, while smaller z-scores represent relatively faster target capture times. Humans performed better in the low 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.

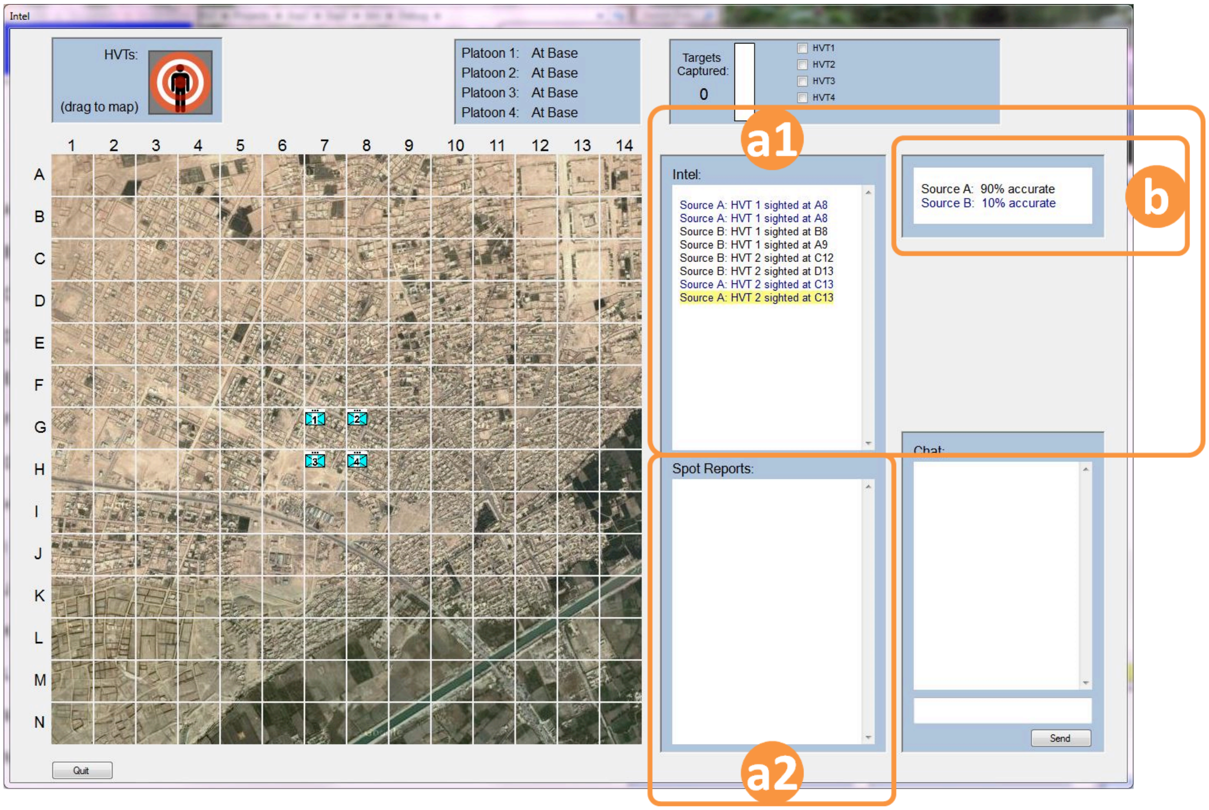


*Figure 6*. Screenshot of the Intel role's display in Experiment 2, with the two experimental manipulations highlighted. 1) In the shared condition, a1 and a2 are shown to both players. In the limited condition, a1 is shown to only the Intel player and a2 is shown to only the OPS player. 2) In the "none" condition, b is hidden. In the congruent condition, the information in b is true. In the incongruent condition, the information in b is false.

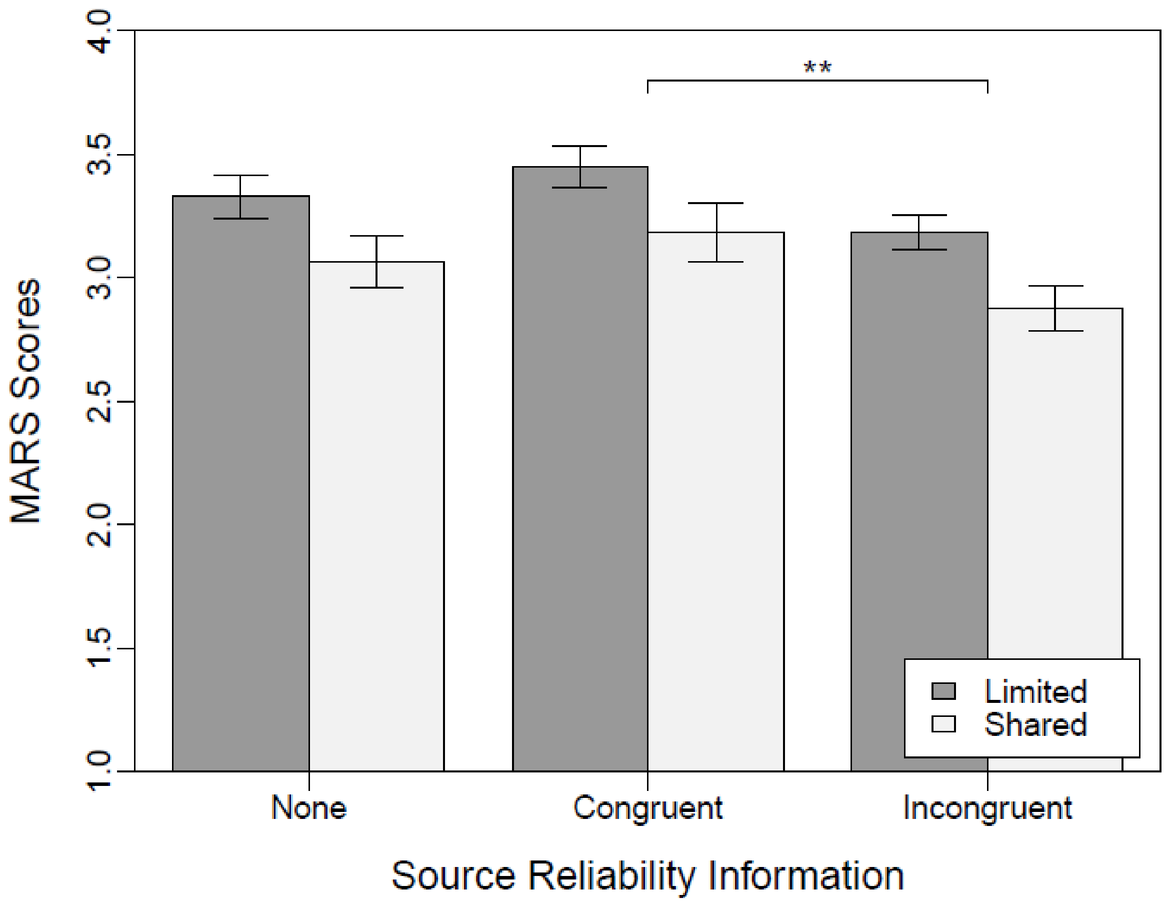


*Figure 7*. Self-reported SA, as measured by MARS scores in Experiment 2. Error bars represent standard errors of the mean.

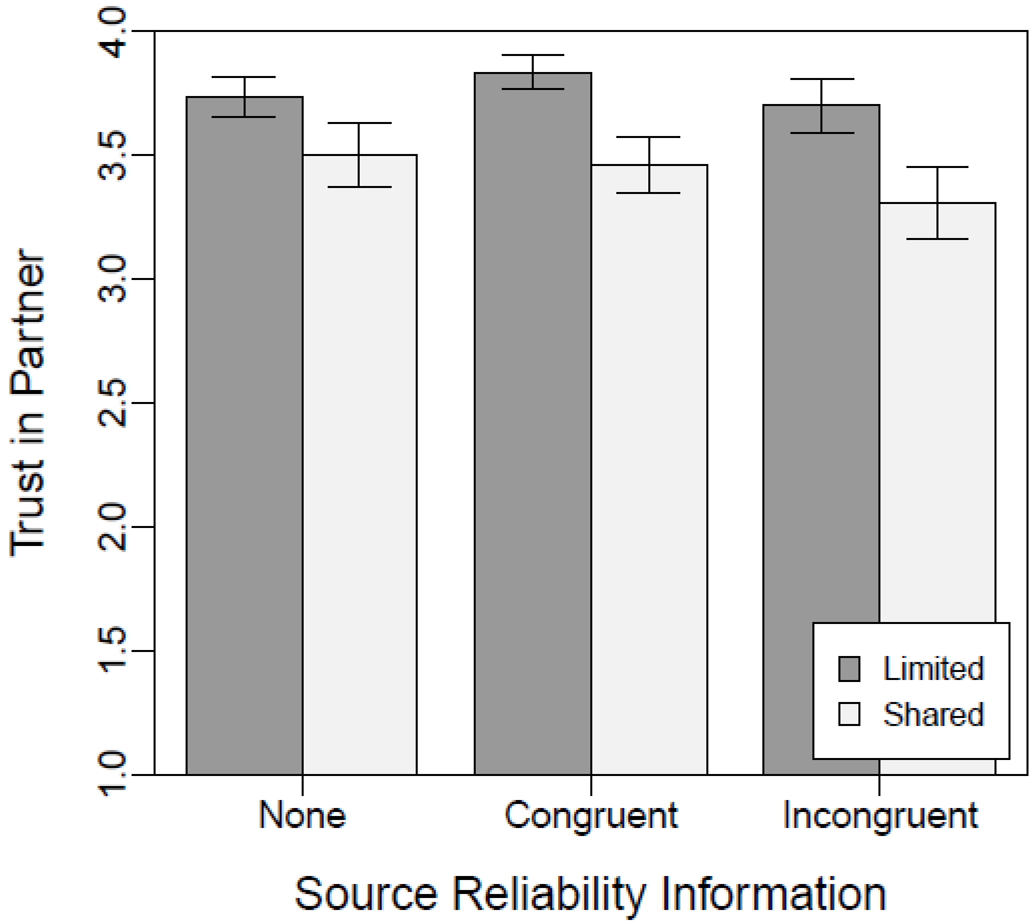


*Figure 8*. Self-reported trust in partner. Error bars represent standard errors of the mean.

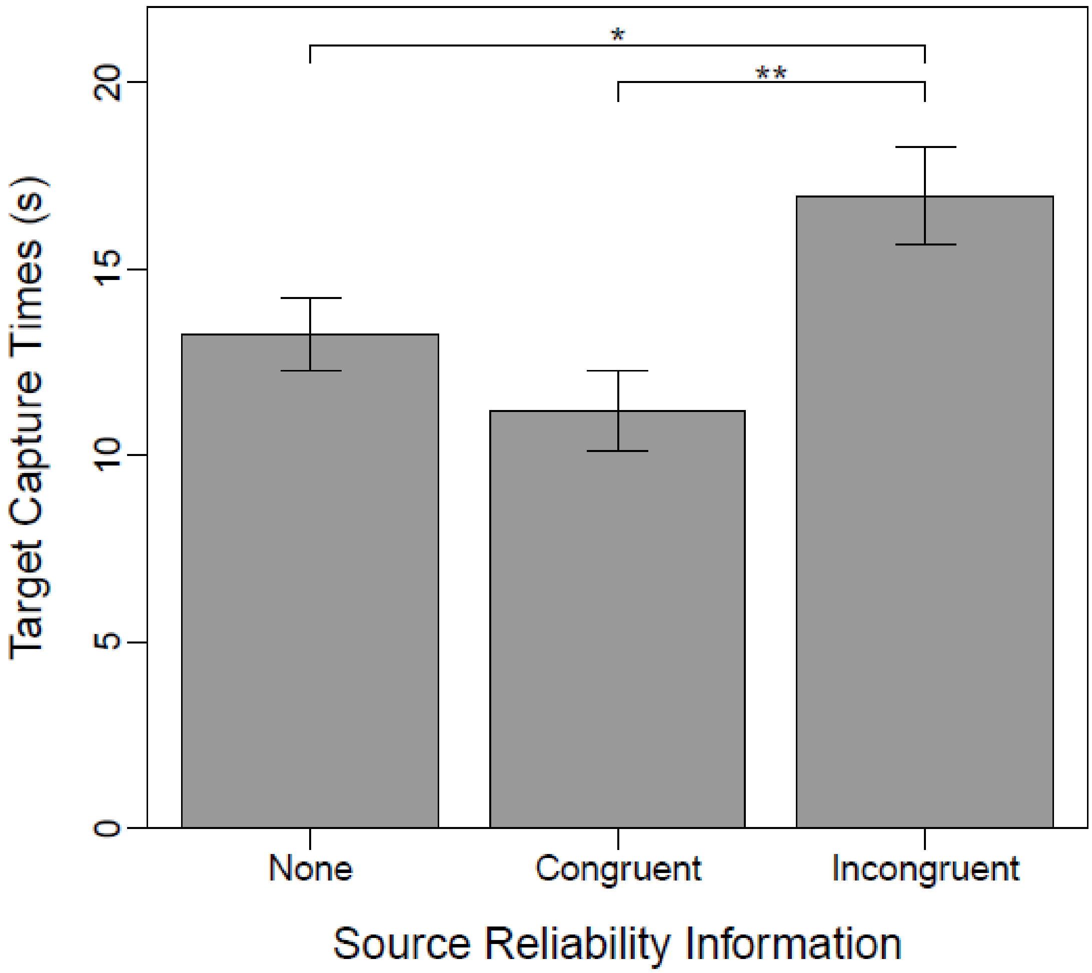


*Figure 9*. Mean target capture times across the three Source Reliability Information conditions in Experiment 2. Error bars represent standard errors of the mean. *$p < 0.05$, **$p < 0.01$

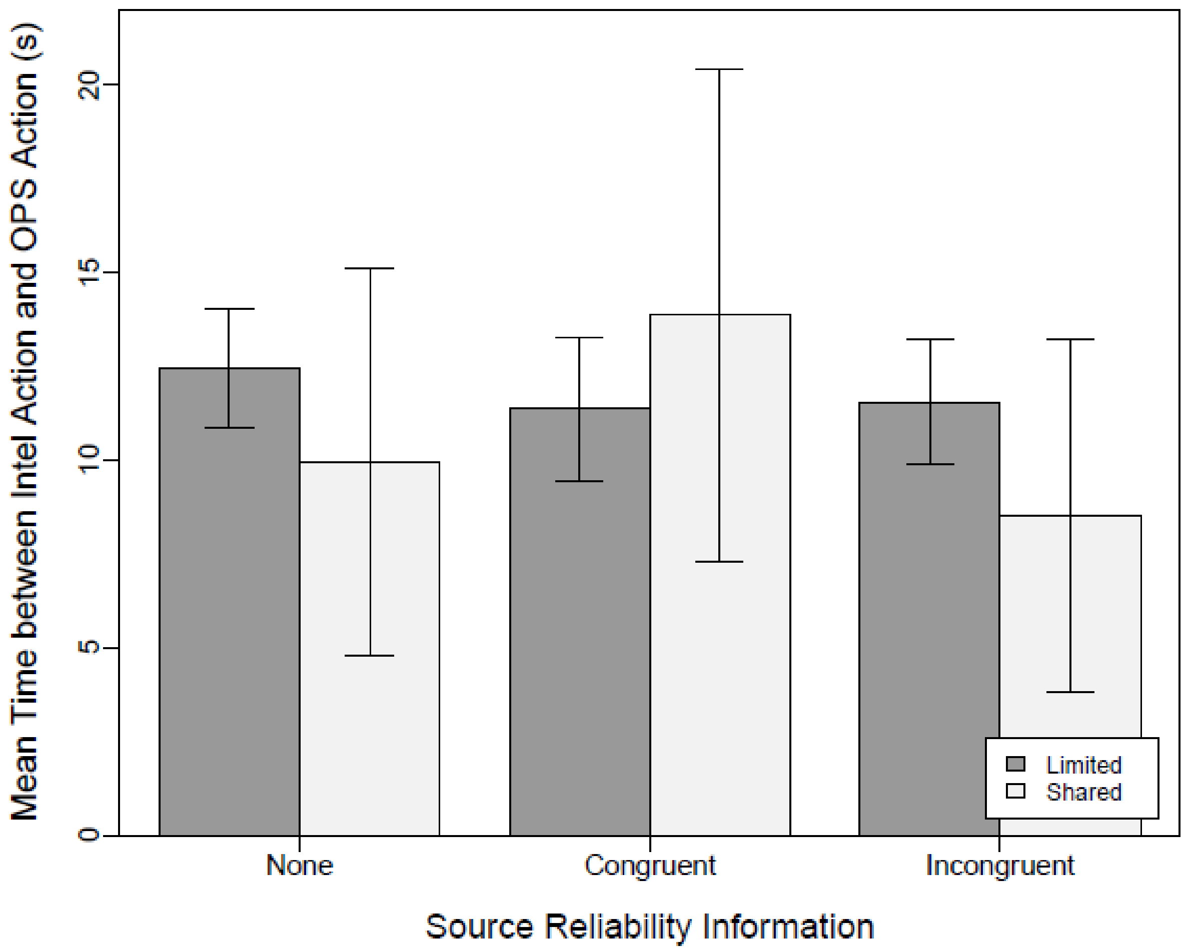


*Figure 10*. Elapsed time after Intel action before OPS action on a given target. Error bars represent standard errors of the mean.

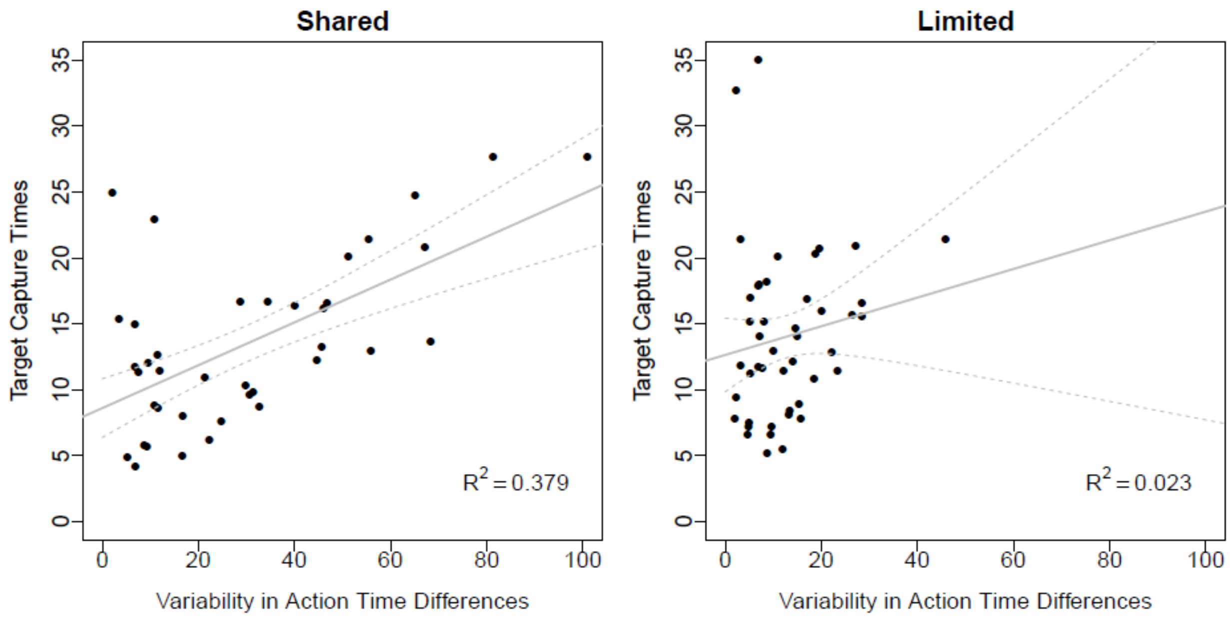


*Figure 11*. Scatterplot showing the relationship between differences in action times and task performance (measured by target capture times). A strong positive correlation exists between these two variables in the Shared conditions.

# Appendix

## Between-Round Questionnaire

Please answer the following questions about the mission you just completed. Your answers to these questions are important in helping us evaluate the effectiveness of this exercise. Check the response that best applies to your experience.

**Please enter your ID number:***

The first four questions deal with your ability to detect and understand important cues present during the mission.

**Please rate your ability to identify critical cues relevant to your task in this mission.***

- ◯ very easy - able to identify all cues
- ◯ fairly easy - could identify most cues
- ◯ somewhat difficult - many cues hard to identify
- ◯ very difficult - had substantial problems identifying most cues

**How well did you understand what was going on during the mission?***

- ◯ very well - fully understood the situation as it unfolded
- ◯ fairly well - understood most aspects of the situation
- ◯ somewhat poorly - had difficulty understanding much of the situation
- ◯ very poorly - the situation did not make sense to me

**How well could you predict what was about to occur next in the mission?***

- ◯ very well - could predict with accuracy what was about to occur
- ◯ fairly well - could make accurate predictions most of the time
- ◯ somewhat poor - misunderstood the situation much of the time
- ◯ very poor - unable to predict what was about to occur

**How aware were you of how to best achieve your goals during this mission?***

- ◯ very aware - knew how to achieve goals at all times
- ◯ fairly aware - knew most of the time how to achieve mission goals
- ◯ somewhat unaware - was not aware of how to achieve some goals
- ◯ very unaware - generally unaware of how to achieve goals

The next four questions ask how difficult it was for you to detect and understand important cues present during the mission.

**How difficult - in terms of mental effort required - was it for you to identify or detect critical cues relevant to your task in the mission?***

- ◯ very easy - could identify relevant cues with little effort
- ◯ fairly easy - could identify relevant cues, but some effort required
- ◯ somewhat difficult - some effort was required to identify most cues
- ◯ very difficult - substantial effort required to identify relevant cues

**How difficult - in terms of mental effort - was it to understand what was going on during the mission?***

- ◯ very easy - understood what was going on with little effort
- ◯ fairly easy - understood events with only moderate effort
- ◯ somewhat difficult - hard to comprehend some aspects of situation
- ◯ very difficult - hard to understand most or all aspects of situation

**How difficult - in terms of mental effort - was it to predict what was about to happen during the mission?***

- ◯ very easy - little or no effort needed
- ◯ fairly easy - moderate effort required
- ◯ somewhat difficult - many projections required substantial effort
- ◯ very difficult - substantial effort required on most or all projections

**How difficult - in terms of mental effort - was it to decide on how to best achieve your goals during this mission?***

- ◯ very easy - little or no effort needed
- ◯ fairly easy - moderate effort required
- ◯ somewhat difficult - substantial effort needed on some decisions
- ◯ very difficult - most or all decisions required substantial effort

The next three questions ask about your teammate in the mission.

**How aware were you of your teammate's activities during this mission?***

- ◯ very aware – knew teammate's activities at all times
- ◯ fairly aware – knew teammate's activities some of the time
- ◯ somewhat unaware – was not aware of some of the teammate's activities
- ◯ very unaware – generally unaware of teammate's activities

**How difficult – in terms of mental effort – was it to share decision making responsibilities with your teammate during this mission?***

- ◯ very easy – could share decision making with little effort
- ◯ fairly easy – could share decision making, but with moderate effort
- ◯ somewhat difficult – effort was required to share decision making
- ◯ very difficult – sharing decision making required substantial effort

**How difficult – in terms of mental effort – was it to effectively communicate with your teammate while completing the tasks in this mission?***

- ◯ very easy – could effectively communicate with little effort
- ◯ fairly easy – could effectively communicate, but with moderate effort
- ◯ somewhat difficult – effort was required to effectively communicate
- ◯ very difficult – effectively communicating required substantial effort

The next three questions ask how much you trust the other player and information sources present during the mission.

**How much did you trust the other player?***

- ◯ Completely
- ◯ Moderately
- ◯ A little
- ◯ Not at all

**How much did you trust Source A?***

- ◯ Completely
- ◯ Moderately
- ◯ A little
- ◯ Not at all

**How much did you trust Source B?***

- ◯ Completely
- ◯ Moderately
- ◯ A little
- ◯ Not at all