11TH ICCRTS

COALITION COMMAND AND CONTROL IN THE NETWORKED ERA

Title of Paper

# The Impact of Instant Messaging on Team Performance, Subjective Workload, and Situation Awareness in Tactical Command and Control

Topics
C2 Experimentation
Cognitive Domain Issues

*Benjamin A. Knott*

General Dynamics, Advanced Information Systems
Wright-Patterson Air Force Base, OH 45433, USA
Benjamin.Knott@wpafb.af.mil

*Robert S. Bolia*
*W. Todd Nelson*
*Scott M. Galster*

Air Force Research Laboratory
Wright-Patterson Air Force Base, OH 45433, USA
Robert.Bolia@wpafb.af.mil
Todd.Nelson@wpafb.af.mil
Scott.Galster@wpafb.af.mil

Point of Contact:

Benjamin A. Knott
General Dynamics Advanced Information Systems
2255 H Street
Area B, Bldg 33, Rm 337
Wright-Patterson AFB, OH 45433-7022
937-255-0866
benjamin.knott@wpafb.af.mil

ABSTRACT

An experiment examined the impact of instant messaging (chat) on team performance, mental workload, and situation awareness in a tactical command and control scenario. This scenario required a team of two weapons directors (WDs) to communicate with each other, their strike packages, and two refueling tankers, to coordinate attacks, defend friendly airspace, and refuel fighter aircraft. Participants conducted the scenario under three communication conditions (voice-only, chat-only, or voice & chat combined). In addition, team workload was manipulated by varying the number of fighter assets available (4 or 8) and the number of enemy targets present in the simulated battlespace (4 or 6). Analyses of team performance scores and subjective measures indicated that chat-only communication resulted in relatively poor team performance, increased mental workload, increased team process demands, reduced situation awareness, and lesser satisfaction with performance. There were no significant differences between the voice-only and the combined (voice and chat) conditions. These results are discussed in terms of their implications for employing chat messaging and other collaboration technologies to support distributed team communication, shared awareness, and decision making in command and control environments.

# 1 Introduction

Network-centric warfare is a model of military operations grounded in the notion that an increasingly dense network of sensors and shooters will promote shared situation awareness at all levels of the command chain and lead to self-synchronized action (Alberts & Hayes, 2003; Alberts, Garstka, & Stein, 2000). These claims suggest the need for collaboration between decision makers at the strategic, operational, and tactical levels of war. There must be a means, for example, of sharing awareness. The network itself is only the architecture. The common operational picture must be provided to fallible human operators in a form that they can comprehend and make use of, and there must be a means by which they can explicitly contribute to this common picture.

Unfortunately, little thought has been given to the design of these technologies in the context of military operations. Their proliferation elsewhere has made enterprise-level collaboration technology suites relatively inexpensive, however, and as a result they have been applied liberally to operational and tactical command and control environments. Usually this is done by analogy, without an analysis of the work domain, the cognitive requirements of the operators, or the appropriateness or usability of the collaboration tools for the tasks they are required to support in the military environment (Scott, Cummings, Graeber, Nelson, & Bolia, 2006).

One of the most frequently used collaboration tools is chat, which allows groups of users on a network to maintain synchronous communication using a text-based format. This modality affords numerous purported advantages, including the ability to sustain multiple conversations simultaneously via different "chat rooms," which is difficult if not impossible using traditional voice communications. Chat also has the potential to relieve the reliance on memory that is often encountered in radio communication, as well as the dependence on the relatively noise-free channels required to preserve intelligibility. The counterarguments suggest that chat requires additional manual control and visual attention relative to voice, and may divert significant resources from the operator's primary task of maintaining and sharing situation awareness, often accomplished through the continuous observation of a geospatial representation of the battlespace and the mouse-driven querying of items on the display. Although a line of reasoning can be developed for both hypotheses, little empirical research has been done to support either.

The purpose of the study described herein was to investigate the effects of chat on the performance, situation awareness, and workload of tactical command and control teams. The context in which the experiment was conducted involved a team of two weapons directors (WD) controlling two combat air patrols in an offensive counter air mission. Each WD was required to communicate with the other, with his own fighter assets, and with air refueling assets in order to successfully complete the mission. This communication could be accomplished using voice, chat, or a combination of the two modalities.

The performance data from this investigation have been reported elsewhere (Knott, Bolia, Nelson, & Galster, 2006). The results suggested a performance decrement associated with the use of chat as the sole

communications modality. This was exhibited by a decrease in the number of successfully prosecuted enemy air targets and an increase in the average prosecution time, as well as an increase in encroachment by enemy fighters into friendly airspace. Due to space limitations and the sheer volume of data collected, only objective measures of mission performance were described in the earlier paper. The purpose of the present paper is to illustrate the effects of chat on subjective measures of operator situation awareness and individual and team mental workload.

# 2 Method

## 2.1 The Distributed Dynamic Decision Making (DDD) Simulator

The DDD[1] was employed to create a set of tactical air battle management scenarios conveyed to participants through a tactical display. The tactical display exhibited the movement of entities within the battle space and provided information about the entities such as speed, heading, weapons and sensor ranges, fuel, and weapons status. The DDD simulator is a reconfigurable simulation designed to support research and training efforts. It allows for rapid development of human-in-the-loop, distributed, multi-person command and control simulations.

## 2.2 The Scenario

A tactical air battle management (TABM) scenario was developed for this experiment and is constructed around a team of two weapons directors (WDs) responsible for the coordination of offensive counter-air, defensive counter-air, and air refuelling operations. The WDs coordinate operations by communicating with each other, their fighter assets, and two refuelling tankers to intercept threats and resupply assets as needed. The scenario was presented to WDs via the DDD tactical display. The tactical display represented the area of operations with friendly assets and enemy targets shown as unique symbols. The display afforded a real-time picture of the battle space from which WDs were able to monitor and direct simulated air operations.

Figure 1 is an example of the WD's tactical display for the TABM scenario. The red and blue symbols represent friendly fighter assets and are labelled with the platform type (e.g., F-18,) and a callsign for each asset. In addition, there are two tankers for aerial refuelling, an air base with two reserve fighter assets, and four infantry units on the ground. The tankers and the air base can refuel the fighter assets and restock their weaponry. Each WD could choose to launch their reserve fighter from the air base as needed. Within the battlespace are Green, Yellow, and Red engagement zones. These demarcations represent different operational areas. The Green zone is the 'kill zone,' the Yellow zone represents friendly airspace, and the Red zone represents the friendly region containing ground assets.

The scenario is a 10-minute simulated counter air operation in which enemy targets enter the green zone and immediately begin moving towards friendly territory. Enemy forces have the ability to attack and destroy all fighter assets, tankers, the air base and the infantry.

Each fighter begins the scenario with weapons resources adequate to complete two attacks on hostile targets, and with a randomly assigned quantity of fuel. The WD's task is to choose appropriate asset-target pairings given the available resources for each asset, communicate the asset-target pairing decisions to friendly fighter assets, and prioritize and coordinate weapons resupply and aerial refuelling with the tankers.

[1] (See http://www.aptima.com/hsi_products_asim.php)

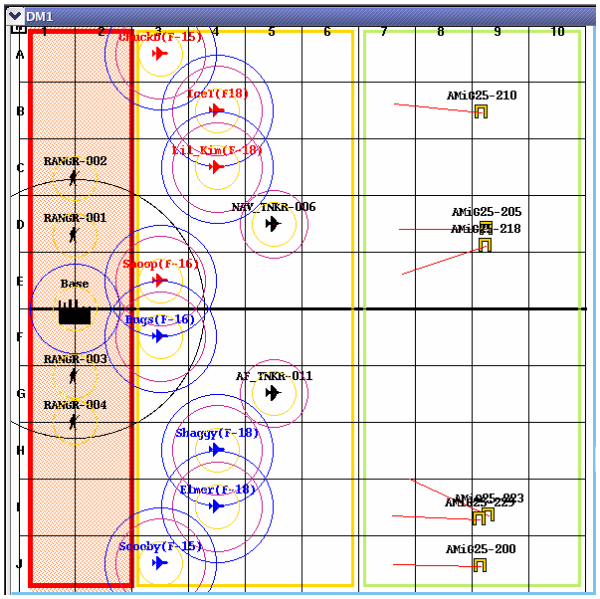


*Figure 1. The WD's tactical situation display. The Red and Blue WD assets are colour-coded and labelled with a callsign. The blue, purple, and yellow rings represent a fighter asset's sensor range, weapons range, and vulnerability to attack, respectively. The enemy targets, labelled as MiG-25s or Su-27s, enter the simulation from the right of the display.*

The two WDs manage separate assets and geographic areas of responsibility (AOR). There are two prescribed AORs, a northern and a southern "lane," and their division was indicated on the tactical display by a solid black horizontal line. Assets are colour-coded such that the WD responsible for the northern lane (the "Red WD") controlled red assets, while the WD responsible for the southern lane (the "Blue WD") controlled blue assets. Although each WD's fighters operate primarily within his or her AOR, fighter assets were able to cross AOR boundaries if necessary to provide assistance by engaging a target for which the other WD's assets had insufficient resources.

The WDs' primary duties include relaying tactical information to their assets, directing assets to intercept hostile targets, and coordinating aerial resupply between assets and tankers. To do this effectively, WDs must understand the capabilities and limitations of their operational environment. Within the simulation, three classes of friendly fighter assets and two classes of hostile targets were employed. Fighters have different fuel capacities and therefore different ranges. All assets have weaponry adequate for two attacks before they must be resupplied. Enemy targets were differentiated by their on-screen representation and their speed of movement. The majority of enemy targets in each scenario were represented by a yellow, inverted "U." These targets, identified as "MiGs," were slightly slower than WD fighter assets and could be pursued and intercepted from behind. The second type of enemy target, identified as "Su-27s," was represented by a red, inverted "V." Such targets were slightly faster than WD fighter assets, rendering pursuit ineffectual, and therefore required frontal interception by fighter assets. Each time a MiG was intercepted and destroyed, a new one would enter the airspace to replace it from the right side of the display. Thus, the number of targets present throughout scenario was deliberately controlled.

The WDs in this scenario are members of multiple teams. The WDs communicate directives to friendly assets through Strike Operators (Red Strike and Blue Strike) and a Tanker Operator. The two Strike Operators play the role of multiple fighter pilots and manoeuvre assets via the DDD interface as directed by their WD. The Tanker Operator manoeuvres the two tankers (an Air Force Tanker and a Navy Tanker) to refuel and resupply assets as directed by the WDs. In this experiment, Strike and Tanker Operators were highly practiced confederates trained to expertise in the role of the strike fighter and tanker operators. As such, their performance is related to, but is not the focus of, team performance in this experiment. Instead, the primary focus is the WDs' task performance.

The team in this scenario included five individuals: Red WD, Blue WD, Red Strike, Blue Strike and the Tanker Operator. The two WDs have all decision making responsibility and direct and manage all air combat operations, coordinate the team, and act on information gleaned from their tactical display and from communication with the other operators. The Tanker Operator and Strike Operators, on the other hand, execute directives from the WDs and also provide the WDs with status updates on their assets, such as fuel and weapons levels.

The WDs' tactical displays provide a global picture of the battle space, including all allied and enemy entities. The Strike Operators are able to see all friendly air or ground assets, but see enemy aircraft only when they come within the limited range of their platform's sensors (represented by a blue ring). Thus, they have limited awareness of the tactical picture and must rely on the WDs to vector them to targets.

The experiment took place in a 9.75 m × 6.5 m room with two WDs on one side of the room and the Strike and Tanker Operators on the other side, facing the opposite direction. Each team member had two 17-inch flat-panel displays. The left display contained radio controls and several chat rooms. ModIOS® Voice Communicator was used for simulated network radio communication. Each of the WD's radios comprised three communication frequencies for speaking with their Strike Operator, the Tanker Operator, and their WD partner. Similarly, Windows Messenger™ was used to create three chat rooms so WDs could communicate via text messages with their Strike Operator, the Tanker Operator, and their WD partner. The right display was used to present the DDD tactical display. WDs were not afforded direct control of the DDD. Rather, they used it to monitor the battle and then used the communications software to issue directives to the Strike and Tanker Operators. The Strike Operators and Tanker Operators used the DDD interface to operate the strike assets and tankers and to retrieve information about their assets. Radios were operated with a footswitch for the Strike and Tanker Operators, and with a mouse for the WDs. Participants wore headsets throughout the experiment and white-noise was generated in the lab at approximately 75 dBA during all trials. The purpose of the white-noise was to simulate the noise of an AWACS or JSTARS platform, and to prevent participants from communicating with each other except by the means provided.

## 2.3 Participants

Six males and 6 females between the ages of 20 and 25 yrs. (Median = 21) were paid to participate in the experiment in six teams of two individuals. There was one all-male team, one all-female team, and 4 teams comprised of one male and one female. All participants reported normal or corrected-to-normal vision in both eyes. Nine of the participants were undergraduate university students and three were graduate students. One participant had military experience (8 months in the Army).

Two teams of three confederates were used for the WD training sessions and during the experiment. The confederates were highly trained individuals who controlled the friendly fighters and the tankers as directed by the WDs and where not the focus of study. Each confederate team was comprised of one Red and one Blue Strike Operator and one Tanker Operator. The confederates conducted 7 hours of training on the operation of assets and the communication protocols prior to the experiment.

## 2.4 Procedure

Prior to the experiment, all participants (WDs) completed a 3-hour training session in which they were trained on the scenario, the radio software, and the chat software, and then completed eight 10-minute practice trials. The trainer informed participants that the purpose of the study was to evaluate how teams used communication technology to work together and that they would be playing a computer game that required teamwork to meet the game's objectives. In addition, WDs were trained on and practiced communication brevity for both voice and text communications. Brevity training was critical to minimize irrelevant, unnecessarily lengthy and/or confusing verbal or text statements. Brevity was also employed to simulate structured military communication. Examples of typical commands for both voice and chat are shown in Table 1. Participants were instructed and trained on using abbreviations in chat messages to reduce the amount of typing required.

*Table 1. Examples of voice and chat brevity communication*

| Participant | Voice command examples | Chat commands examples |
|---|---|---|
| Blue WD | "Elmer move to H7" | "EL move to H7" |
| Red WD | "ChuckD intercept at J3" | "CD Int, J3" |
| Blue WD | "Scooby fuel status?" | "SC fuel stat?" |
| Red WD | "Snoop refuel at Air Force Tanker" | "SN refuel, AFT" |
| Blue WD | "Bugs refuel at the Base" | "BU refuel, Base" |

Participants were also trained on the specific objectives and rules of the mission, displayed in Tables 2 and 3, and were instructed that the performance of the team would be measured for each trial based on how well they met their objectives and followed the rules.

*Table 2. Mission objectives for the TABM task.*

| | **Mission Objectives** |
|---|---|
| 1 | Destroy as many hostile aircraft as quickly as possible. |
| 2 | Do not allow hostile aircraft to enter friendly territory (Yellow and Red Zones). |
| 3 | Protect the Air Base and the infantry from enemy attack. |
| 4 | Protect the Air Force and Navy Tankers from enemy attack. |
| 5 | Keep as many fighters airborne for as long as possible |

*Table 3. Refueling rules for the TABM task.*

| | **Mission rules** |
|---|---|
| 1 | Navy fighters (the F-18s) must be refuelled at the Navy Tanker. Air force fighters (F-15s & F-16s) must be refuelled at the Air Force Tanker. |
| 2 | Do not refuel at the Base unless an airborne refuelling is not possible. |

One to two days after training was completed, WDs returned for the experimental session. In this session they experienced one practice trial to re-acclimate them to the task and then completed twelve 10-minute experimental trials. After each trial, participants completed several subjective instruments designed to assess, among other things, situation awareness, mental workload, the demands of team processes, and satisfaction with task performance. Participants were given one 10-minute rest period after they had completed half of the trials. All major simulation events (e.g., the occurrence and outcome of attacks, refuelling events, etc.) were recorded in data logs for later analysis. In addition, video of the all team members and all voice and chat communications were recorded.

## 2.5 Experimental Design

There were three levels of Communications Modality (voice-only, chat-only, voice & chat combined), two levels of the number of fighter assets the WDs had to manage (4 or 8), and two levels of the number of enemy targets present in the battlespace throughout the scenario (4 or 6). These independent variables were combined factorially, yielding a 3 × 2 × 2 within-subjects design. In the voice & chat condition, all members of the team were given the option to communicate using voice, chat, or a combination of both. The variation in the number of assets and targets provided two different workload manipulations. The order of conditions was counterbalanced across trials.

### *2.5.1 Subjective Measures*

Two instruments were used to evaluate perceived task demands. A modified version of the NASA Task Load Index (Hart & Staveland, 1987) was used for ratings of team workload. In the modified version, participants were asked to estimate the workload for the *team* rather than rating their individual workload. For this experiment, only the subscale measures were recorded. The second instrument comprised five subscales that provided ratings of workload unique to team processes. These were used to assess the demands of Communication, Monitoring, Control, Coordination, and Leadership (Table 4). Each sub-scale consisted of three questions which participants rated from 1 (low demand) to 20 (high demand).

*Table 4. The team process workload sub-scales.*

| **Sub-scale** | **Description** |
|---|---|
| Communication Demand | The demands associated with communication between team members. |
| Monitoring Demand | The demands associated with monitoring others during the scenario. |
| Control Demand | The demands associated correcting of others during the scenario. |
| Coordination Demand | The demands of adjusting one's own actions during the scenario. |
| Leadership Demand | The demands associated with providing leadership during the scenario. |

A measure of a the WD's overall situation awareness (SA) was obtained using a question derived from the Situation Assessment Rating Technique (SART) (Taylor, 1990), which asked participants to rate their overall SA on a scale from 1 (low SA) to 7 (high SA).

WDs were also asked to rate, on a 7-point scale, satisfaction with their own performance (3 questions) and satisfaction with the performance of the team (3 questions) after each trial. The three questions for each scale were averaged to arrive at satisfaction scores for individual and team performance.

### *2.5.2 Team Performance*

A single measure of team performance on the task was calculated by averaging 1) the percentage of enemy targets that were allowed to penetrate friendly airspace, 2) the percentage of high value assets destroyed (the air base, infantry units, and tanker aircraft), and 3) the percentage of fighter assets that were lost when they were allowed to run out of fuel. The average was then subtracted from 100, resulting in a team performance score in which 100 indicated optimal team performance in accordance with the mission objectives. More detail on team performance measures for this task, and the rationale for the team performance score, is provided by Knott, Bolia, Nelson, & Galster (2006).

# 3 Results

## 3.1 Team Performance

A measure of team performance, described above, was calculated for each trial. An analysis of the team performance scores revealed a main effect for Communication Modality, $F(2, 10) = 13.59$, $p < .05$. The post-hoc analysis indicated that the source of the effect was relatively low team performance scores for the chat conditions compared to either the voice-only or combined voice and chat conditions (Table 5). An interaction between the numbers of Assets and Targets was also obtained, $F(1, 5) = 38.14$, $p < .05$, as depicted in Figure 2. Post-hoc comparisons revealed that team performance was highest with 4 assets and 4 targets and lowest when there were only 4 assets to pair with 6 targets. The number of targets did not have an effect on performance in the 8 asset conditions as evident in the figure.

*Table 5. Mean team performance score as a function of communication modality.*

| Communication | Mean | Std Error |
|---|---|---|
| Voice | 82.36 | 1.68 |
| Chat & Voice | 79.33 | 1.96 |
| Chat | 73.82 | 2.21 |

*Figure 2. Mean team performance score as a function of the number of fighter assets and the number of targets in the scenario. Error bars represent one standard error of the mean in each direction.*

## 3.2 Team Workload

Participants were asked to rate the mental workload of their *team* on each of the six subscales of the NASA Task Load Index (Hart & Staveland, 1987). The unweighted subscale ratings for the two WDs on each team were averaged together and the mean scores for each subscale subjected to a 6 (Subscale) × 2 (Number of Targets) × 2 (Number of Assets) × 3 (Communication Modality) repeated measures analysis of variance, which revealed main effects of Subscale, $F(5, 25) = 30.21$, $p < .05$; Communication Modality, $F(2, 10) = 6.91$, $p < .05$; and a Subscale × Number of Assets interaction, $F(5, 25) = 2.62$, $p < .05$. All other sources of variance lacked statistical significance.

The main effect of Communication Modality, depicted in Figure 3, was further explored by means of post-hoc paired *t*-tests, which determined that workload ratings were higher under the chat condition than under either of the other conditions involving voice communications. This finding is consistent with the main effect of Communication Modality on team performance shown above in Table 5.

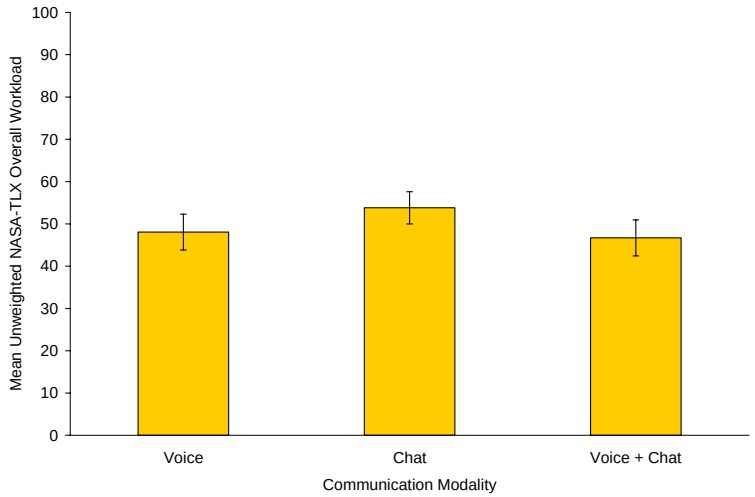


*Figure 3. Mean unweighted NASA-TLX ratings as a function of Communication Modality. Error bars indicate one standard error of the mean in each direction.*

The 2-way interaction is plotted in Figure 4. The source of the interaction was found to be the increase in perceived team effort as the number of assets controlled by the WD increased. Analogous results were not obtained for any of the other subscales. Inspection of the figure also provides a useful picture of the workload profile of the scenario employed, which engendered fairly high mental and temporal demands, required a high level of effort, and induced a reasonable amount of frustration.

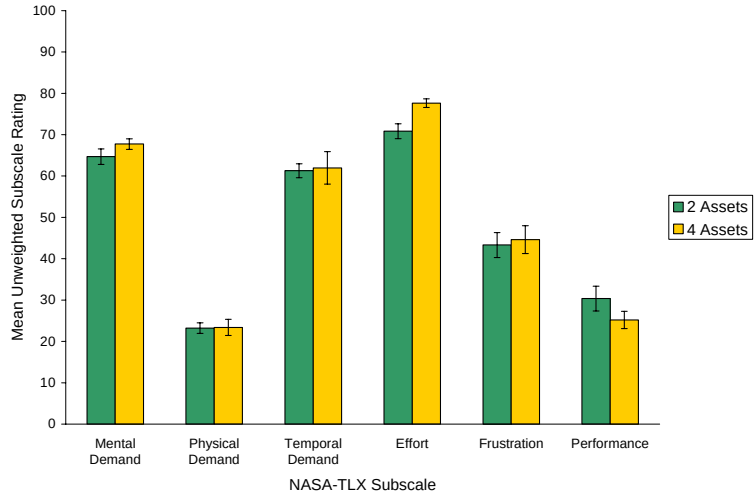


*Figure 4. Mean unweighted NASA-TLX subscale ratings as a function of subscale and the number of targets in the scenario. Error bars represent one standard error of the mean in each direction.*

## 3.3 Team Process Demands

After completing the six subscales of the NASA-TLX, participants were asked to rate the demands related to team processes on the five subscales in Table 4. Ratings for each of the 12 participants were submitted to a 5 (Subscale) × 2 (Number of Targets) × 2 (Number of Assets) × 3 (Communication Modality) repeated measures analysis of variance, which revealed main effects of Subscale, $F(4, 44) = 4.57$, $p < .05$, and Communication Modality, $F(2, 22) = 5.08$, $p < .05$. All other sources of variance lacked statistical significance.

The main effect of Communication Modality, depicted in Figure 5, was further explored by means of post-hoc paired *t*-tests, which determined that team process demands were rated higher under the chat condition

than under either of the conditions involving voice communications. There was no difference between the voice-only condition and the condition using voice and chat. These findings are consistent with both main effects of Communication Modality on team performance (Table 5), and NASA TLX ratings (Figure 3).

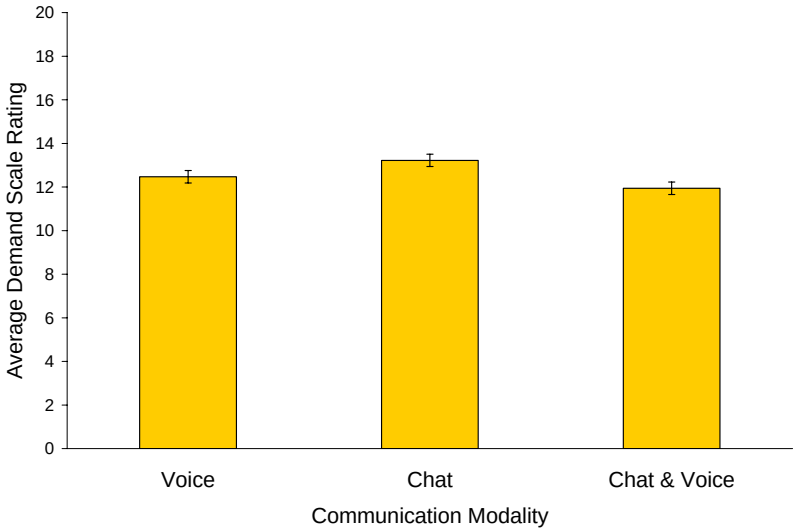


*Figure 5. Mean team process demand ratings as a function of communication modality. Error bars represent one standard error of the mean in each direction.*

The main effect of Subscale is shown in Figure 6 which indicates relatively high Leadership, Communications, and Monitoring demands, when compared to demands for Coordination and Control. Note however, that ratings were above the midpoint on all subscales suggesting that the scenario in use elicited considerable demands on these team processes overall. Figure 6 provides a useful profile for team process demands on the scenario employed. Specifically, the scenario engendered relatively high Leadership, Communications, and Monitoring demands with lesser demands for Coordination and Control. The team process subscales together with the subscales illustrated in Figure 4, serve as a workload profile for the current scenario and provide a basis of comparison for future manipulations or for related task scenarios.

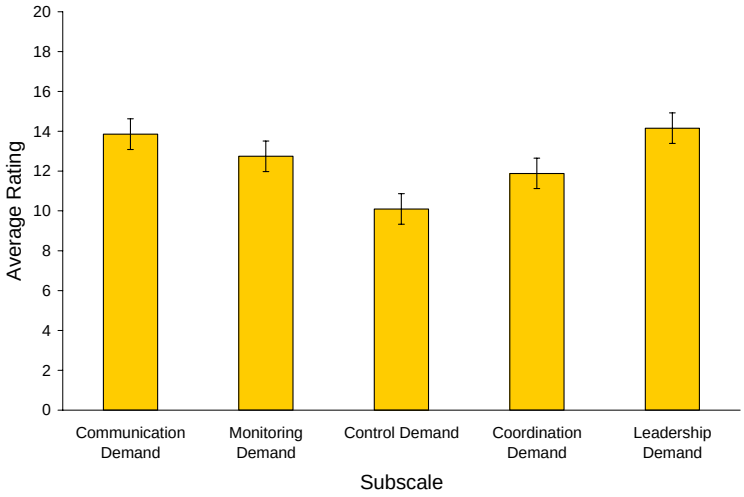


*Figure 6. Mean demand ratings as a function of team process subscale. Error bars represent one standard error of the mean in each direction.*

## 3.4 Situation Awareness

A measure of a WD's overall situation awareness (SA) was obtained using a question derived from the Situation Assessment Rating Technique (SART; Taylor, 1990), which asked participants to rate their overall SA on a scale from 1 (low SA) to 7 (high SA). Responses from 11 participants – one WD neglected to complete the survey – were submitted to a 2 (Number of Targets) × 2 (Number of Assets) × 3 (Communication Modality) repeated measures ANOVA. The only significant source of variance revealed by the analysis was a main effect of Communication Modality, $F(2, 20) = 7.12$, $p < .05$, illustrated in Figure 7. Post-hoc inspection identified the source of this effect as the reduction in overall SA associated with the use of chat compared to either the voice or combined conditions.

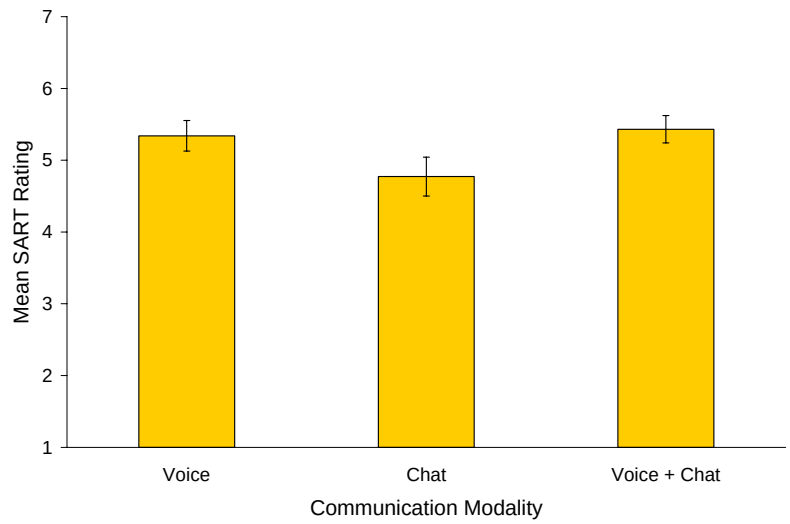


*Figure 7. Mean overall SA ratings as a function of communication modality. Error bars indicate one standard error of the mean in each direction.*

## 3.5 Satisfaction with Task Performance

Team and individual satisfaction ratings for the 12 WDs were submitted to a 2 (Number of Targets) × 2 (Number of Assets) × 3 (Communication Modality) repeated measures ANOVA. Main effects for Communication Modality were obtained for satisfaction ratings of both individual, $F(2,22) = 6.32$, $p < .05$, and team performance, $F(2,22) = 5.34$, $p < .05$. Consistent with the previous measures, post-hoc tests revealed that satisfaction with individual and team performance were significantly lower for the chat condition compared to both of the voice conditions (Table 6 and Table 7). It is noteworthy that the satisfaction ratings of individual and team performance were highly correlated $r = .86$, $p < 05$.

*Table 6. Mean rating of satisfaction with individual performance as a function of Communication Modality.*

| Communication | Mean | Std Error |
|---|---|---|
| Chat & Voice | 5.70 | 0.16 |
| Voice | 5.50 | 0.15 |
| Chat | 5.00 | 0.21 |

*Table 7. Mean rating of satisfaction with team performance as a function of Communication Modality.*

| Communication | Mean | Std Error |
|---|---|---|
| Chat & Voice | 5.94 | 0.17 |
| Voice | 5.85 | 0.14 |
| Chat | 5.32 | 0.18 |

# 4 Discussion

The purpose of this research was to evaluate the performance and perceived workload effects of using a chat application either alone or in combination with voice communication for communication and collaboration in a tactical command and control task. Teams of two weapons directors were trained on a TABM scenario which they performed under three Communication Modality conditions (chat-only, voice-only, or chat & voice combined). Workload was varied by manipulating the number of targets or the number of assets in each trial. The communication modality results indicated that chat-only led to relatively poor team performance, increased mental workload, increased team process demands, reduced situation awareness, and lesser satisfaction with performance on the task. Interestingly, no significant differences were obtained between the voice-only and the combined (voice and chat) condition.

The finding that chat, when used as the sole communication modality, engenders a group performance decrement is generally consistent with the existing body of literature on text-based computer mediated communication (CMC). Literature reviews of studies that compared face-to-face (FTF) communication, in which voice is the primary modality, with text-based CMC groups, consistently show that CMC is associated with increased decision time, greater difficulty in reaching group consensus, and decreased individual satisfaction with group processes. While CMC groups are more likely to have greater equality of participation in team processes, they also result in less participation by group members overall (Wainfan & Davis, 2004; Baltes, Dickson, Sherman, Bauer & LaGanke, 2002). Notably, the literature also shows that although it takes groups longer to reach a decision when using text messaging, there is no loss in decision quality for CMC compared to FTF groups, when the groups have unlimited time to reach decisions (Wainfan & Davis, 2004; Baltes, et al. 2002). This is, of course, a condition not typically afforded by the time pressures of TABM. However, this finding does suggest that if the specific factors responsible for the slowing of group decision making processes for CMC can be identified and addressed, then perhaps CMC, such as chat, may be employed as an effective distributed collaboration tool.

Several explanations may be offered for the observed disadvantage of using chat alone for communication in TABM. First, the monitoring of chat rooms consumed visual resources that might have been better directed toward the situation display. Monitoring chat rooms likely disrupted the WD's sampling of the tactical situation, reducing his or her awareness. This explanation is partially supported by the analysis of SART responses which indicated reduced situation awareness for the chat-only condition. Second, text entry required the Strike and Tanker Operators to take their hands off the mouse, which was their primary control device, preventing them in some cases from performing their "flight" tasks as efficiently. Third, the abbreviations that WDs and operators used for chat communication (see Table 1) may have caused additional "translation delays" as transmitters and receivers had to think about what to type, and translate abbreviations prior to taking action. This possibility will be examined in future studies. Finally, chat communication might have required more time for the WDs and their operators to type and read directives when compared to speaking and listening over the radio. The communication delays imposed by typing chat messages may have lead to relatively poorer coordination and less efficient performance on time-critical aspects of the TABM task. Several previous studies have noted that text, compared to voice communications, is slower and more effortful (Siegel, Dubrovsky, Kiesler, & McGuire, 1986; Bordia, 1997), resulting in less frequent communication and extended decision times.

The failure to find differences on any measures between the voice-only and combined (voice and chat) communication would appear to be inconsistent with recent research reported by Cummings (2004), in which operators were required to monitor and retarget cruise missiles from a missile control station while monitoring chat as a secondary task. The author proposed that chat, as a secondary task, interfered with performance on the primary task of missile retargeting because chat messages constituted an interruption that shifted the operator's attention away from the primary task. However, the present study evaluated the role of chat when communication between group members was the operators' primary activity, rather than a secondary event. In this context, having chat for supplementary communication did not appear to have a negative impact on either performance or perceived workload.

This conclusion, however, should be considered with caution as additional analysis is needed better understand how teams exploited chat as secondary communication channel. It is possible that there are no differences between voice-only and the combined communication condition because, in the combined condition, which allowed participants to use voice and chat in any manner they chose, operators simply did not use chat very often. Another possibility is that high-performing teams more effectively adapted their technology usage in a manner appropriate for certain sub-tasks. For instance, operators may have used voice for more dynamic and time-critical communications (such as guiding an asset to a target), and used text when temporal demands were relaxed (such as asset status updates). This hypothesis is based on comments gathered from the experiment debriefing questionnaire in which many participants indicated that although they preferred voice communication overall, chat was effective is certain situations.

To more clearly understand how chat was used in the combined condition within this TABM scenario, an analysis of both the frequency and content of each communication channel (voice or chat) is underway. This analysis will include an evaluation of the relative frequency with which participants used chat and voice to communicate with each team member, and, more importantly, an examination of the content of chat messages compared to voice transmissions. A content analysis will involve transcribing all voice communication and then classifying each voice utterance and text message according to a predefined taxonomy. Analysis of the message classifications will provide a rich source of information for how groups used the voice and chat tools to communicate within this context.

# 5 Conclusions and Future Directions

Reviews of the literature on collaboration technology consistently emphasise the importance of task characteristics when interpreting the findings concerning the use of CMC. Specifically, collaboration technology has been shown to affect outcomes differently depending on the characteristics of the task performed (Wainfan & Davis, 2004; Hiltz, Johnson, & Turoff, 1986). Consideration of the type of task that a team is engaged in is important because different types of collaboration technology are likely to be better suited to different types of tasks (Bolstad & Endsley, 2005). For instance chat messaging has been shown to be effective for brainstorming and idea generation (Valacich, Dennis, & Connolly, 1994; Gallupe, Bastianutti, & Cooper, 1991; Dennis, & Valacich, 1993), while FTF or video conferencing is more effective for negotiating or resolving disputes (Wainfan & Davis, 2004; Arunachalam & Dilla, 1995). One widely accepted taxonomy of task types is the McGrath circumplex model (1984), which posits four task categories: 1) generation; 2) choice; 3) negotiation; and 4) execution. The majority of research on collaboration technology and CMC is concerned with the first three of these categories (Baltes et al., 2002), specifically, those that involve generating ideas, choosing among alternatives, or resolving conflicts. Command and control activities, such as TABM, fall into the fourth category (execution), in which CMC is used to coordinate a team's action in response to the actions of an opposing team. The study presented here represents an important step in addressing an area of research that has received relatively little attention: the effect of collaboration technology on teams engaged in the execution of military operations.

While the focus of this study was on the use of chat within the TABM context, the goal of research on collaboration technology for command and control should ultimately be to understand how a variety of different tools may be integrated into this domain to best support the operator engaged in the task. There have been several attempts at devising methods to determine what type, or types, of collaboration technology are most suitable to different situations, collaborative processes, or tasks, based on previous research. Wainfan & Davis (2004) outlined a strategy for selecting the best tools for distributed teams to collaborate based on the characteristics of the technology, team composition, and other mediating factors. For example, they recommend the use of chat rooms for a brainstorming task based on research that shows that individuals are less inhibited and groups generate more ideas when using text messaging compared to FTF brainstorming sessions. Bolstad & Endsley (2005) also present a collaboration taxonomy that can be used as a guide to selecting the best collaboration tools for a particular task based on collaboration and technology characteristics, information requirements, and task type. However, Bolstad & Endsley (2005) argue that although there are a variety of general purpose collaboration tools available, some situations demand domain-specific tools when commercially available technology does not adequately address the needs of the users within a specific domain. A domain specific tool is one that is tailored to the needs of the

individuals in a well defined task domain. One conclusion from the present study is that the teams could benefit from domain-specific tools that are shaped by the unique demands of the TABM task and the individual roles within that task.

Although the general purpose chat application used in this study resulted in relatively poor performance, compared to voice, a simple customization of this tool to the TABM scenario could address the speed of communication issue (i.e. typing delay) discussed above. Within the scenario, the communication between group members is both highly structured and constrained by the task. For instance, in a typical WD-to-Strike Operator transmission, the WD would identify an asset, the action to be taken, and a location (e.g., "Elmer, intercept at B8"). In a highly constrained communication environment such as this it may be possible to increase the speed of communication for chat by partially automating the generation of text messages. This can be achieved by employing custom menus or buttons associated with callsigns, actions, and commands, so that operators can construct messages with a few mouse clicks rather than typing a complete text string. In this example, a domain-specific chat tool, which is a simple modification of an existing commercial application, has the potential to greatly improve communication effectiveness by reducing the time to transmit a message, allowing for quicker and more consistent dialog, giving operators more time to attend to their situation display.

Future research will examine additional domain-specific tools to meet the unique information and interaction needs of the particular task and environment. This investigation need not be limited to text and voice communication tools. The spatial nature of the TABM task lends itself well to tools that support communication through graphics or images, such as digital whiteboard or file sharing applications. A custom whiteboard application may allow structured commands involving complex movement sequences to be conveyed graphically in a manner that is easily translated into actions by the receiver of the graphical message. Future studies will endeavor to examine the influence of chat messaging on command and control tasks, while also exploring the application of novel, custom, or domain-specific tools as a way to optimize or tune the collaboration environment to the a specific task domain. Ultimately, the goal is to understand the interplay between various communication modalities and collaboration technologies, including communication by voice, text, video, and images, to promote communication effectiveness, shared awareness, and decision quality in command and control domains.

# 6 References


Alberts, D. S., Garstka, J. I., & Stein, F. P. (2000). *Network-centric warfare: Developing and leveraging information superiority*. Washington: Command & Control Research Program.

Alberts, D. S., & Hayes, R. E. (2003). *Power to the edge: Command… control… in the information age.* Washington: Command & Control Research Program.

Arunachalam, V. & Dilla, W. N. (1995). Judgment accuracy and outcomes in negotiation: A causal modeling analysis of decision-aiding effects. *Organizational Behavior and Human Decision Processes*, 61, 289–304.

Baltes, B.B., Dickson, M.W., Sherman, M.P., Bauer, C.C., & LaGanke, J.S. (2002). Computer-mediated communication and group decision making: A meta-analysis. *Organizational Behavior and Human Decision Processes, 87*, 156-179.

Bolia, R. S., Nelson, W. T., Vidulich, M. A., Simpson, B. D., & Brungart, D. S. (2005). Communications research for command and control: Human-machine interface technologies supporting effective air battle management. *Proceedings of the 10th International Command and Control Research and Technology Symposium.* Washington: Command and Control Research Program.

Bolstad, C.A. & Endsley, M.R. (2005). Choosing team collaboration tools: Lessons from disaster recovery efforts. *Ergonomics in Design*, 13, 7-14.

Bordia, P. (1997). Face-to-face versus computer-mediated communication: A synthesis of the experimental literature. *The Journal of Business Communication*, 34, 99-121.

Cummings, M. L. (2004). The need for command and control instant message adaptive interfaces: Lessons learned from tactical tomahawk human-in-the-loop simulations. *CyberPsychology & Behavior, 7*, 653-661.

Dennis, A. R. & Valacich, J. S. (1993). Computer brainstorms: More heads are better than one. *Journal of Applied Psychology*, 78, 531–537.

Dubrovsky, V. J., Kiesler, S., & Sethna, B. N. (1991). The equalization phenomenon: Status effects in computer-mediated and face-to-face decision-making groups. *Human-Computer Interaction*, 6, 119-146.

Gallupe, R. B., Bastianutti, L. M., & Cooper, W. H. (1991). Unlocking brainstorms. *Journal of Applied Psychology*, 76, 137–142.

Hart, S.G. & Staveland, L.E. (1987). 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).

Hiltz, S. R., Johnson, K., & Turoff, M. (1986). Experiments in group decision making: Communication process and outcome in face-to-face versus computerized conferences. *Human Communication Research*, 13, 225-252.

Knott, B. A., Bolia, R. S., Nelson, W. T., Galster, S. M. (2006). Effects of collaboration technology on the performance of tactical air battle management teams. *Proceedings of the Symposium on Human Factors Issues in Network-Centric Warfare*, Sydney, Australia.

McGrath, J. E. (1984). *Groups: Interaction and Performance*, Englewood Cliffs, NJ: Prentice Hall.

Reid, F., Ball, L., Morley, A., & Evans, J. (1997). Styles of group discussion in computer-mediated decision-making. *British Journal of Social Psychology*, 36, 241–262.

Scott, S. D., Cummings, M. L., Graeber, D. A., Nelson, W. T., & Bolia, R. S. (2006). Collaboration technology in military team operations: Lessons learned from the corporate domain. *Proceedings of the 2006 Command and Control Research and Technology Symposium.* Washington: Command and Control Research Program.

Siegel, J., Dubrovsky, V., Kiesler, S., & McGuire, T. W. (1986). Group processes in computer-mediated communication. *Organizational Behavior and Human Decision Processes*, 37, 157-186.

Taylor, R.M. (1990). Situational Awareness Rating Technique (SART): The development of a tool for aircrew systems design. In: *AGARD Conference Proceedings, Situational Awareness in Aerospace Operations*. Aerospace Medical Panel Symposium, Copenhagen.

Valacich, J. S., Dennis, A. R., & Connolly, T. (1994). Idea generation in computer-based groups: A new ending to an old story. *Organizational Behavior and Human Decision Processes*, 57, 448–467.

Wainfan, L. & Davis, P. (2004). *Challenges in Virtual Collaboration: Videoconferencing, Audioconferencing, and Computer-Mediated Communications*, Santa Monica, CA: RAND Corporation, TK5105.6.W35 2005.

Weisband, S. P. (1992). Group discussion and first advocacy effects in computer-mediated and face-to-face decision making groups. *Organizational Behavior and Human Decision Processes*, 53, 352-380.