Effects of Collaboration Technology on the Performance of Tactical Air Battle Management Teams

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

Robert S. Bolia
Air Force Research Laboratory
Wright-Patterson Air Force Base, OH 45433, USA
Robert.Bolia@wpafb.af.mil

W. Todd Nelson
Air Force Research Laboratory
Wright-Patterson Air Force Base, OH 45433, USA
Todd.Nelson@wpafb.af.mil

Scott M. Galster
Air Force Research Laboratory
Wright-Patterson Air Force Base, OH 45433, USA
Scott.Galster@wpafb.af.mil

Abstract

Collaboration technologies offer the potential to enhance communication effectiveness, shared awareness, and decision quality in command and control (C2) application domains. The purpose of this experiment is to evaluate the impact of one of these technologies – Instant Messaging (Chat) – on team performance efficiency in a simulated human-in-the-loop air defence C2 scenario. An air battle management task required a team of two weapons directors (WDs) to communicate with each other, their fighter assets, and two refuelling tankers, to coordinate offensive counter-air, defensive counter-air, and air refuelling. The study compares the impact of three communications conditions (Voice-only, Chat-only, Voice and Chat) under varying levels of workload by manipulating the number of enemy targets (4 or 6) present in the battle space and the number of fighters (4 or 8) that the two WDs must manage. Team performance efficiency is assessed in terms of the proportion of friendly assets and enemy targets destroyed, the number of enemy target incursions into friendly airspace and the time required to eliminate the enemy targets. The results are discussed with respect to the efficacy of text–based collaboration technologies for supporting distributed team decision making in future C2 environments.

1 Introduction

Air battle management refers to the command and control (C2) of air combat operations, and involves the management of assets engaged in strike, interdiction, and close support operations, as well as offensive and defensive counter air, air refuelling, and air mobility missions. At the operational level this is accomplished in the Combat Operations Division of an Air and Space Operations Center (AOC), which is responsible for the execution of the daily Air Tasking Order and the management of the time-sensitive targeting process. At the tactical level, air battle management is handled by weapons directors (WDs) distributed throughout the area of operations in a number of airborne platforms, including the USAF E-3 Airborne Warning and Control System (AWACS) and E-8 Joint Surveillance Target Attack Radar (JSTARS), or the USN E-2C Hawkeye and various fixed or mobile surface assets such as the USN Tactical Air Control Centers (TACC) and USAF Control and Reporting Centers (CRC).[1]

The difference in the missions of these diverse tactical air battle management (TABM) systems is tied to their sensor capabilities. AWACS and JSTARS, for example, are able to provide radar coverage over vast sections of enemy territory, and hence are useful for the control of counter-air, deep strike, or interdiction missions. The Hawkeye has the capability to control similar missions, but because it is carrier-launched and has limited range it is optimized for counter ship and fleet defence operations. CRCs, on the other hand, are land based and tied to short range land-based radars, and find their utility in the control of air operations near the forward edge of the battle area, while TACCs are ship-based and are designed to control air operations in the littoral.

[1] Note that this is not an exhaustive enumeration of tactical air battle management platforms. All four US armed services have a range of options for tactical air battle management, and the armed forces of other nations further diversify the array. See Williams (1997) and Armistead (2002) for more information on the airborne platforms.

The multiplicity of TABM assets notwithstanding, WDs across platforms perform similar tasks with similar displays and controls. Their primary display is a geospatial representation of the area of operations, overlaid with political and operational demarcations, terrain data, and symbology representing friendly and enemy assets. This provides them with a real-time picture of the dynamic battlespace, within the limits of sensor coverage and identification and tracking accuracy. Their primary input modality is a keyboard augmented by some form of pointing device, usually a trackball. These are used, for example, to determine geometric relationships among the displayed entities, such as bearing and range, to manipulate track information, and to communicate text or other visual information to colleagues on the platform.

The primary task of a WD is the control of air assets. This includes the vectoring of aircraft to intercept air and ground targets, the control of planned strike missions, the sharing of the tactical picture with other platforms, and the coordination of air refuelling and combat search and rescue. It is important to note that this task is not performed in isolation. The WD should be viewed as a member of a number of teams, or teams of teams. He is a member, for example, of the weapons team on his platform (for example, a cohort of WDs led by a senior director); of the team comprised by the platform as a whole (for example, the AWACS mission crew, which includes the weapons and surveillance teams); of the mission team that includes his platform as well as the assets under his control (for example, a strike package with associated refuelling, combat air patrol, and suppression of enemy air defence); and so on. This aspect of the WD's role is significant because it stresses the importance of collaboration for the achievement of tactical and operational goals.

In current TABM environments, collaboration is accomplished primarily by means of voice communication, with each WD monitoring and transmitting on multiple radio and intercom channels. This may not be optimal, however, owing to the propensity for communications overload and the possible reduction in speech intelligibility due to ambient platform noise (Bolia, Nelson, Vidulich, Simpson, & Brungart, 2005). The proliferation of collaboration technologies (Nelson, Bolia, Vidulich, & Langhorne, 2004) such as "chat" or instant messaging has suggested an alternative to traditional voice communication.

The use of "chat rooms" in lieu of radio or intercom networks offers a number of prospective advantages to air battle managers. First, multiple chat rooms can be maintained simultaneously and, unlike voice channels, are not masked by speech or environmental noise. Second, they provide a text-based record of the communication on each channel, reducing the memory requirements of operators and the requirement to request retransmission of missed information. Finally, they provide immediate visual feedback about the source of the communication, which is often absent in radio transmission. There are, on the other hand, potential drawbacks to the use of this modality. One is the fact that it introduces additional input requirements that may detract from the operator's usual use of his hands – voice communication is typically hands-free, since a footswitch is used to activate the channel. The other is that the operator's visual attention may be diverted from his geospatial display, which is his primary means of maintaining situation awareness. To date, these possibilities have not been critically evaluated.

There is a growing literature on the effect of collaboration technologies (CT) and computer-mediated communication (CMC) on team performance. However, most of the research is concerned with comparing groups collaborating face-to-face with distributed groups communicating with computer-mediated collaboration tools. In addition, many published CT studies (Baltes, Dickson, Sherman, Bauer, & LaGanke, 2002) focus on business applications involving tasks that require a group to arrive at consensus on a decision or to generate ideas. These studies may not generalize well to a C2 environment in which the task is to execute a mission. The purpose of the present investigation is to compare the performance of teams using radio voice communication with teams employing text messaging as either a primary or supplementary means of communication in a high-workload military C2 environment.

2 Method

2.1 The DDD Simulator

The Distributed Dynamic Decision-Making (DDD)[2] software is a tool for creating human-in-the-loop, distributed, multi-person simulations. The DDD was employed to create a set of TABM simulations 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.

2.2 The Scenario

The TABM scenario developed for this experiment 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 by communicating with each other, their fighter assets, and two refuelling tankers. The scenario was presented to WDs via the DDD tactical display. The tactical display represented the area of operations with friendly and enemy assets 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 and callsign for each asset. In addition, there are two tankers for aerial refuelling, an air base, and four infantry units on the ground. The tankers and the air base can refuel the fighter assets and restock their weaponry. 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.

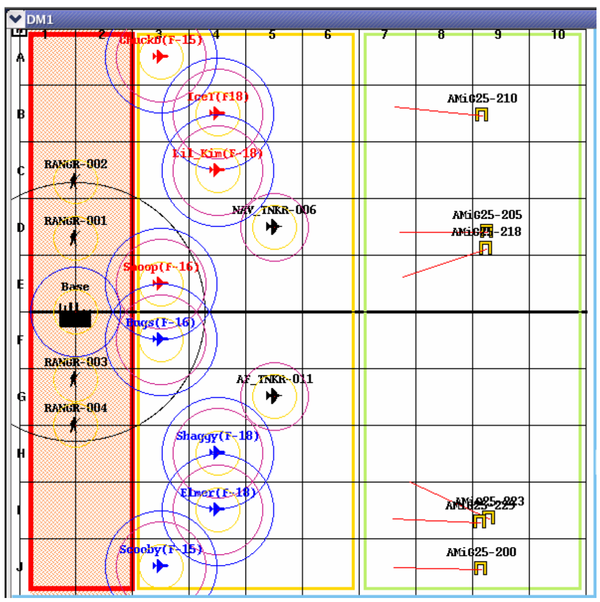

Figure 1. The WD's tactical display. The Red and Blue WD assets are colour-coded and labelled with a callsign. The enemy targets, labelled as MiGs, enter the simulation from the right of the display.

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.

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

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.

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 refuelling 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 "Su27s," were 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 fighter and tanker crews. 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. 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 experimental confederates 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. The right 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 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 confederates, and with a mouse for 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 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, the chat software, and 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 communications brevity for both voice and text communications. Brevity training was critical to minimize irrelevant, unnecessarily lengthy or confusing verbal or text statements. This was also done to simulate structured military commutation.

Participants were also trained on the specific objectives and rules of the mission, displayed in Table 1. Objectives are listed in the order of priority as presented to participants. Participants 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 1. Mission objectives and rules and their link to dependent measures.

	Mission Objectives	**Dependent Measures**
1	Destroy as many hostile aircraft as quickly as possible.	• Percentage of Targets Intercepted • Average Time to Intercept
2	Do not allow hostile aircraft to enter friendly territory (Yellow and Red Zones).	• Percentage of Yellow Zone Incursions.
3. 4.	Protect the Air Base and the infantry from enemy attack. Protect the Air Force and Navy Tankers from enemy attack.	• Percentage of High Value Asset Losses
5	Keep as many fighters airborne for as long as possible	• Percentage of Fighter Assets Losses due to: a. enemy attack b. out of fuel
	Mission rules	**Dependent Measures**
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.	• Percentage of Refuelling Errors
2	Do not refuel at the Base unless an airborne refuelling is not possible.	• Percentage of Assets refuelled at the base

One to two days after training was completed, WDs returned for the experimental session. In this session they completed 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 and perceived workload. Participants were given one 10-minute rest period upon completion of 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, audio, and chat communications were recorded for each team member.

2.4 Experimental Design

Three levels of communications (Voice-only, Chat-only, Voice & Chat), two levels of fighter assets (4 or 8 fighters), and two levels of enemy targets (4 or 6 targets present in the battle space) were combined factorially, yielding a 3 × 2 × 2 within-subjects design. In the Voice & Chat communication condition, all members of the team were given the option to use Voice, Chat, or some combination of the two to communicate with their team. The variation in the number of assets and targets provided two different workload manipulations. The presentation order of the conditions was counterbalanced across trials.

2.5 Participants

Six teams of two individuals were paid to participate as WDs in the experiment. Overall there were 6 men and 6 women between the ages of 20 and 25 (Median = 21). There was one all-men team, one all-women team, and 4 teams comprised of one man and one woman. 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. Only one participant had military experience (8 months in the Army).

Two teams of three confederates were also paid to participate in the experiment. The confederates were highly trained individuals whose responsibilities were limited to the control of friendly assets as directed by the WDs. They were not thus the focus of study. Each confederate team was comprised of a red and a blue strike operator and a tanker operator. The confederates received seven hours of training on the operation of assets and the communication protocol prior to the experiment.

3 Results

A Communication × Assets × Targets within-subjects ANOVA was conducted for each of the performance measures listed in Table 1. When appropriate, Tukey's HSD procedure was used for post-hoc multiple comparisons. Due to space limitations, analyses of subjective measures, communications patterns, and refuelling errors will be reported elsewhere (Knott, Bolia, Nelson, & Galster, 2006).

3.1 Percentage of Targets Intercepted

Main effects were obtained for Assets (F (1, 5) = 14.19, p <.05) and Communication (F (2, 10) = 12.86, p < .05). When teams had more assets available, the percentage of available targets intercepted was higher (M = 69.3% & 73.1%, for 4 and 8 assets, respectively). Interestingly, the Voice & Chat communication condition resulted in a higher percentage of intercepts than either the Voice-Only or the Chat-Only conditions. The difference between the Voice-Only and the Chat-Only conditions was not significant (Table 2).

Table 2. Percentage of targets intercepted as a function of communications modality.

Communication	Mean	Std Error
Voice & Chat	74.0%	1.3%
Voice	71.0%	1.0%
Chat	68.5%	1.3%

3.2 Average Time to Intercept

The average time to intercept a target was defined as the time elapsed between the appearance of a target in the scenario and its prosecution by a fighter asset. On average, targets were intercepted more quickly when teams had more assets (M= 152.9 and M =140.3 seconds for 4 and 8 assets, respectively), F (1, 5) = 31.66, p < .05, and when there were fewer targets, F (1, 5) = 10.18, p <.05. The interaction of Targets with Communication was also significant, F (2, 10) = 4.23, p < .05. Figure 2 suggests that there were no differences in performance as a function of communication in the 6 targets condition, but the Voice & Chat condition resulted in faster intercepts compared to the Chat-Only and Voice-Only conditions with 4-targets.

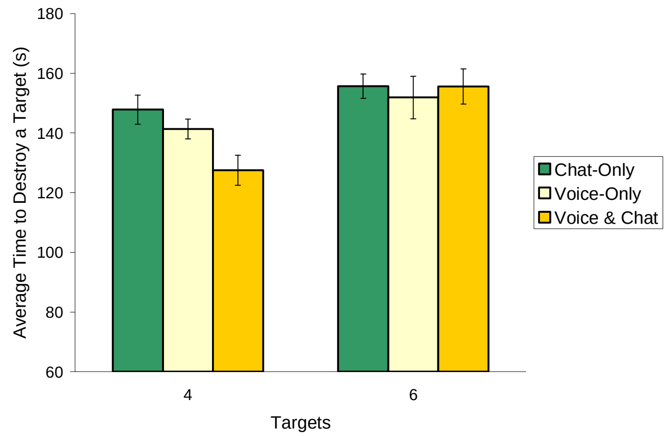

Figure 2. Average intercept time as a function of communications modality and the number of targets in the scenario.

3.3 Percentage of Yellow Zone Incursions

This measure is the percentage of targets that move from the Green to the Yellow Zone before being intercepted. There were main effects for Targets, F (1, 5) = 11.02, $p < .05$, and Communication conditions, F (2, 10) = 6.03, $p < .05$, and a significant interaction for Assets × Targets, F (1, 5) = 55.68, $p < .05$. In addition, there was a significant three-way interaction between Assets, Targets, and Communication, F (2,10) = 7.70, $p < 0.05$. Means and standard errors for each condition are plotted in Figure 3. The figure suggests that the interaction of Assets and Communication differs across Target conditions. Since we are primarily interested in the effects of communication, the simple 2-way interactions of Communication × Assets were compared separately for 4-targets and 6-targets conditions with a Scheffé adjustment for post hoc testing. This analysis revealed that the simple interaction of Communication and Assets was significant for 4-targets, F (2,10) = 8.48, $p < .025$, but was not significant for 6-targets. Multiple comparison analyses indicated that, for 4-targets, there was no difference in performance between 4 or 8-assets per trial for the Voice & Chat or Voice-Only conditions. However, there was a significant difference between 4 or 8-assets for Chat communication, with 8-assets resulting in poorer performance overall. A simple main effect analysis of Assets was also performed separately for the 4-targets and 6-targets conditions. A significant main affect for Assets was obtained for the 6-target conditions, such that 8-assets resulted in better performance on average, compared to 4-assets, F (1, 5) = 24.48, $p < .025$. There was no main effect of Assets in the 4-target conditions. These analyses suggest that the main effect of communication on this measure was due primarily to the 8-asset, 4-target, chat condition which resulted in particularly poor performance compared to the other 4-target conditions. Also of interest is the finding that having more assets improved performance for 6-targets, but had no effect for 4-targets.

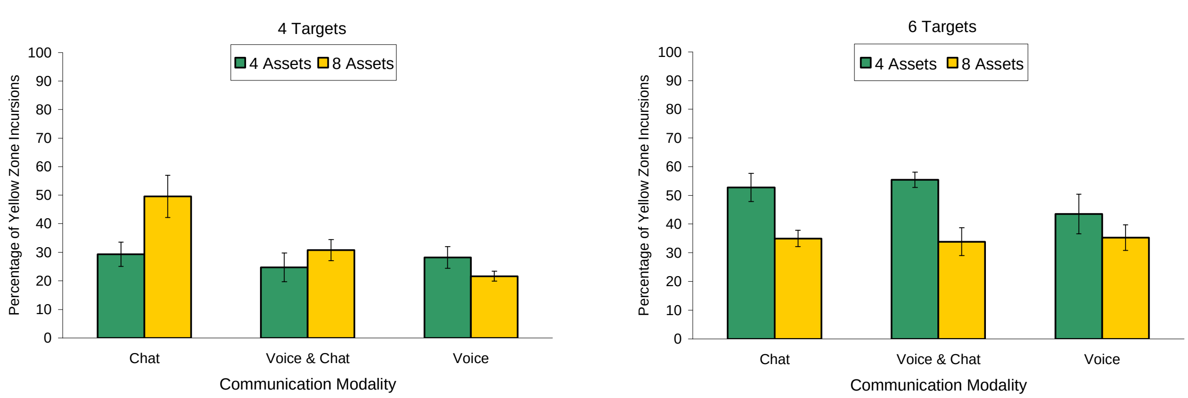

Figure 3. The three-way interaction of Communication, Assets and Targets for the percent of yellow zone incursions.

3.4 Percentage of Fighter Asset Losses

Two analyses were conducted to explore the effect of the experimental manipulations on fighter asset losses. The first is the percentage of asset losses to enemy attack. There were no significant effects for assets lost to enemy attack. The second measure is the percentage of fighter asset losses due to fuel depletion. This measure can be considered a measure of the WD's effectiveness in terms of planning, coordinating, and prioritizing refuelling operations. There were significant effects for Assets, F (1, 5) = 35.24, $p < .05$, and Communication, F (2, 10) = 7.21, $p < .05$ for assets lost due to fuel depletion. Asset losses were, on average, less than one fighter per trial, (M = 6.9%) when the team had 4 assets to manage, while the losses were approximately 1 in 5 when the team had 8 assets to manage (M = 20.8%). The main effect of communication indicated that Chat-Only lead to higher losses than Voice-Only. There was no difference between Voice & Chat and the other communication conditions.

Table 3. Percentage of fighter assets lost due to fuel depletion as a function of communications modality.

Communication	Mean	Std Error
Voice	8.8%	2.2%
Voice & Chat	13.9%	3.2%
Chat	19.0%	3.2%

3.5 Percentage of High–Value Assets Losses

The 'high-value' assets in this scenario are the two tankers, the four infantry units, and the air base. The only significant source of variance disclosed by the analysis was an Assets × Targets interaction, F (1, 5) = 7.50, $p < .05$, displayed in Figure 4. Team performance significantly worsened as the number of targets increased for 4-asset trials, while there was no performance decrement between 6- and 4-targets during the 8-assets trials. As team assets increased, WDs were able to protect high-value assets more effectively in the high target condition.

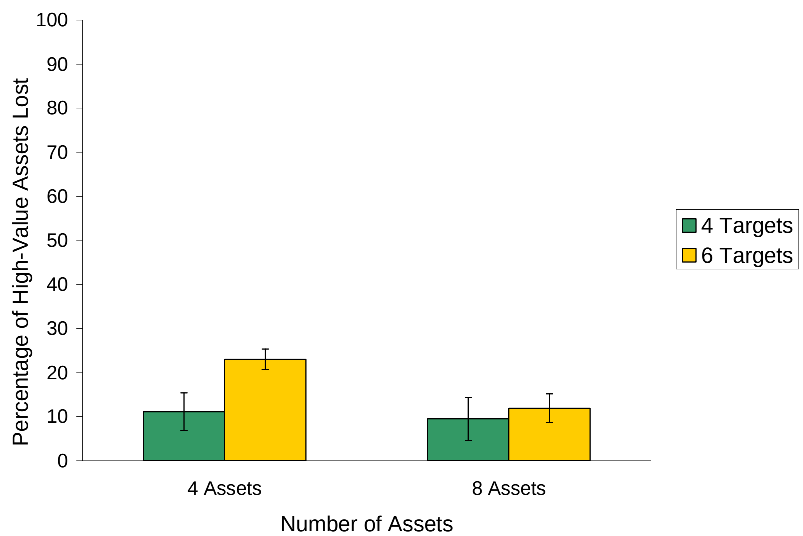

Figure 4. Average high-value asset losses as a function of the number of assets and the number of targets in the scenario.

3.6 Composite Team Performance Score

A composite score was calculated as a measure of overall team performance on the task. The composite score was calculated as 100 minus the average of the following dependent measures: 1) Percentage of Yellow Zone Incursions; 2) Percentage of high value Asset Losses; and 3) Percentage of Fighter Asset Losses due to fuel depletion. This resulted in composite team performance on a scale from 0 to 100. A score of 100 indicated optimal team performance in accordance with the mission objectives. Mean composite scores for each communication condition are displayed in Table 4. An analysis of the team composite 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 4). An interaction between the numbers of Assets and Targets was also obtained, $F(1, 5) = 38.14$, $p < .05$, as depicted in Figure 5. Post-hoc comparisons revealed that team performance was highest with 4 assets and 4 targets in the scenario and lowest when there were 4 assets to pair with 6 targets. Increasing the number of targets from 4 to 6 did not effect performance when the team had 8 assets as evident in the figure.

Table 4. 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 5. Team composite score as a function of the number of assets and the number of targets in the scenario.

4 Discussion

The results of this investigation suggest a performance decrement associated with communication restricted to chat. Teams performed less effectively in the chat-only condition compared to the Voice & Chat condition (though not less effective than Voice alone) in terms of the percentage of targets intercepted, and the average time to intercept targets, albeit only in the 4-target condition for the latter measure. The percentage of asset losses due to fuel depletion is of particular interest because of the time-critical nature of aerial refuelling. On this measure, the Chat-Only condition resulted in more than double the proportion of losses than the Voice-Only condition, but the observed difference between Chat-Only and Voice & Chat was not reliable. While it is somewhat surprising that there were no significant differences between the two voice conditions on target intercepts and time to intercept, the failure to find differences may have been due to the small sample size.

The differences between the Voice-Only and the Voice & Chat conditions were not as clear. The combined (Voice & Chat) condition yielded a significant advantage compared to Voice-Only on the percentage of targets intercepted and the average time to intercept targets in the 4-targets condition. However, there was no difference between the two conditions on asset losses due to fuel depletion. One possible explanation for these findings may be participants' overall preference for voice communication in the Voice & Chat condition. The primary purpose of this condition was to allow team members, including the confederates, to flexibly communicate by voice, chat, or both. As such, participants were not required to use chat during those trials, and it is possible that chat was used relatively infrequently. If so, then the Voice & Chat condition should be viewed as primarily a voice condition with chat as an additional, secondary means of communication. A complete analysis of the communication data will be necessary to characterize the differences in team communication behaviour between two voice conditions. Although not all measures of team performance revealed differences between the Voice-Only and Voice & Chat conditions, the difference in the percentage of targets intercepted provides evidence that chat may have in some respects provided an advantage when used as a communication channel secondary to voice.

The finding that chat employed as the sole communication modality engenders a performance decrement in a TABM task is consistent with the existing body of literature on CT and CMC. For example, Bordia (1997) conducted a review of 18 studies comparing text-based communication with face-to-face (FTF) verbal communication on group performance for decision making and idea generation tasks. This review suggests that while the quality of decisions did not differ between CMC and FTF communication groups, text messaging CMC groups took longer to complete tasks (Hiltz, Johnson & Turoff, 1986; Weisband, 1992; Reid, Ball, Morley, & Evens 1997; Siegel, Dubrovsky, Kiesler, & McGuire, 1986). Bordia concluded that this effect is primarily a technical limitation associated with text messaging. It takes longer to type and read massages than to speak and listen. Thus, if task time is not limited, the effectiveness of CMC and FTF groups is not different. In a more recent comparison of FTF and CMC groups, Baltes et al. (2002) conclude that the use of CMC is associated with decreased group effectiveness, increased time to decision, and decreased individual satisfaction. However, similar to Bordia (1997), Baltes et al. (2002) hypothesized that CMC groups can be as effective as FTF groups when task time is not limited. The TABM scenario in the present study is both communication-intensive and time-limited. WDs must rapidly send messages, either by voice or text, to coordinate air defence and refuelling efficiently. If chat does not support the rapid communication afforded by speech, then it is reasonable to expect that performance on time critical tasks, such as aerial refuelling, would be degraded as a result.

While speed of communication is one possible explanation for the observed performance differences between the two modalities, there are other factors that must be considered. First, the monitoring of chat rooms consumes visual resources that might otherwise be directed toward the situation display. It is possible that monitoring the chat rooms may have disrupted WDs' sampling of the tactical situation, reducing his awareness. Second, text entry in chat required the strike and tanker operators to take their hands off the mouse, which was their primary input control device, and may have prevented them from performing their primary "flight" tasks as efficiently. Third, the abbreviated text participants were instructed to use in this study may have caused additional "translation delays" as they formatted outgoing messages and interpreted incoming ones. Although the use of abbreviations is quite common for text messaging, it is not clear that the training provided in the present study was sufficient for WDs to reach expertise in their use. Finally, text messaging is a relatively impoverished communication channel, lacking the paralinguistic and social cues often present in human voice communication. Future studies will need to address these questions for TABM scenarios. For example, methods to reduce communication delays imposed by typing could be explored using abbreviations and semi-automated messaging; such an approach may be plausible within highly structured communication environments such as TABM.

Overall, this study does not suggest any disadvantage to using chat as a secondary means of communication. This is important because in many military environments, chat is used as a supplementary means of communication. These results are also interesting when compared to those of Cummings (2004), who required operators to monitor and retarget cruise missiles while monitoring chat as a secondary task. Cummings found that chat degraded primary task performance, and posited that this result was due to operators' distraction from the primary task by chat. The present study, on the other hand, evaluated the role of chat when communication is part of WDs' primary task, rather than a secondary event. The results of this experiment indicate that chat, as an additional communications channel, did not negatively impact performance, and in the case of the target intercept measures, provided a slight advantage over voice alone.

The present research represents one of the few empirical studies to examine synchronous text messaging (chat) as a means of both primary and secondary communication. In addition, the majority of CT or CMC research compares text communication with FTF communication. Within a distributed C2 domain it is more appropriate to compare chat with voice communication via radio transmission. This is an important distinction, given that there are numerous problems with mediated voice communications, such as poor voice quality and added cognitive demands associated with managing and monitoring multiple radio channels.

The use of chat as a communication tool is likely to become more pervasive in the coming years within both civilian and military organizations. It is therefore important to understand both the benefits and the deleterious effects of chat on the C2 process. Future studies will need to explore the performance decrements observed to date (e.g., speed of communication, diverted attention, etc.) and generate suggestions to ameliorate those effects. Ultimately, the goal is to understand the interplay between chat and other collaboration technologies, including communication by voice, text, video, and imagery, to promote communication effectiveness, shared awareness, and decision quality in C2 domains.

5 Acknowledgements

The authors gratefully acknowledge the contributions to this research of the following individuals: Becky Brown, April Bennett, and Dan Schwartz, for their role in data collection and management, and training; Brent Miller and Allen Dukes, without whose creative and efficient software support the experiment would not have been run at all; Jim Hyson, for greatly simplifying the compilation and management of the simulation data; Lt Col A. C. Shaw, Chris Best, and Mike Skinner, for useful discussions on scenario design; Capt Dave Miller, Lt Dustin Weeks, for their assistance in the development of the experiment and in data collection; Greg Funke, for technical assistance during data collection and his critical review of this work; Eduardo Salas for his insights into the conduct of team research and discussions of metrics; and Doug Brungart for his review of an earlier draft of this paper.

6 References

Armistead, E. L. (2002). *AWACS and Hawkeyes: The complete history of airborne early warning aircraft*. St. Paul, MN: MBI Publishing.

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.

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.

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.

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., Nelson, W. T., Bolia, R. S., & Galster, S. M. (2006). The impact of instant messaging on team performance, subjective workload, and situation awareness in air battle management. *Proceedings of the 11th International Command and Control Research & Technology Symposium*. Washington: Command and Control Research Program.

Nelson, W. T., Bolia, R. S., Vidulich, M. A., & Langhorne, A. L. (2004). User-centered evaluation of multi-national communications and collaborative technologies in a network-centric air battle management environment. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting* (pp. 731-735). Santa Monica, CA: Human Factors and Ergonomics Society.

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

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.

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.

Williams, G. K. (1997). AWACS and JSTARS. In J. Neufeld, G. M. Watson, & D. Chenoweth (Eds.), *Technology and the Air Force: A retrospective assessment* (pp. 267-287). Washington: Air Force History and Museums Program.