15th ICCRTS
"The Evolution of C2"

Development and Evaluation of the Multi-Modal Communication Management Suite

Topic 5: Experimentation and Analysis

Victor S. Finomore, Jr.
Air Force Research Laboratory
2610 Seventh Street B 441
Wright-Patterson AFB OH, 45433
937-904-7123
victor.finomore@wpafb.af.mil

Dianne K. Popik
Air Force Research Laboratory
2610 Seventh Street B 441
Wright-Patterson AFB OH, 45433
937-255-4381
dianne.popik@wpafb.af.mil

Brian D. Simpson
Air Force Research Laboratory
2610 Seventh Street B 441
Wright-Patterson AFB OH, 45433
937-255-4463
Brian.simpson@wpafb.af.mil

Douglas S. Brungart
Walter Reed Medical Center
6900 Georgia Ave NW
Washington DC, 20307
douglas.brungart@us.army.mil

Courtney E. Castle
Oak Ridge Institute for Science and Education
2610 Seventh Street B 441
Wright-Patterson AFB OH, 45433
937-904-7905
courtney.castle@wpafb.af.mil

Ron C. Dallman
Ball Aerospace
2610 Seventh Street B 441
Wright-Patterson AFB OH, 45433
937-255-3980
ron.dallman@wpafb.af.mil

Abstract

Command and Control operators rely heavily on radio communications to efficiently plan, direct, coordinate, and control assets to successfully complete their mission. This communication intensive environment imposes a high degree of workload on operators thus resulting in failures of detection or comprehension of communication. Multi-Modal Communication (MMC) is a net-centric communication management suite that captures, visually displays, records, and archives radio communication thus allowing instant access to past communications as well as an increase in the intelligibility of current communication. This study examined the performance associated with monitoring multiple communication channels with access to different communication tools. In all experimental conditions, operators participated in a 10-minute communication-monitoring task in which they monitored and responded to the occurrence of critical phrases. Communication performance was analyzed in regards to message detection, response accuracy, and response time. Data showed that MMC can provide a balance between the speed of radio listening and the accuracy and data-capturing capabilities of chat displays. MMC can be a beneficial tool to Command and Control operators in its ability to increase intelligibly while providing a persistent, searchable visual display of voice communication, which reduces workload while helping maintain situation awareness.

Introduction

The success of a mission relies heavily on the ability of the Command and Control (C2) operator to effectively communicate with their assets. These operators are responsible for the real-time management of multiple distributed teammates, who are all accountable for different aspects of the mission. The C2 operator monitors communications from each asset to build a picture of the dynamic battlefield and directs key players based on its evolution.

This communication-intensive environment imposes a high workload on the operators, who typically monitor and transmit on eight or more simultaneous channels (Bolia, 2003). A major issue is that multiple operators can speak at the same time, thus reducing the intelligibility of essential messages. An additional problem is the transient nature of radio transmissions. The recipient has one chance to extract crucial information or is required to request that the information be repeated. Not only is missed information a problem during the mission, but the operators must take detailed notes to ensure the correct information is relayed to the appropriate persons. This added step to relay the information to the correct asset opens up the possibility for error and may place the mission in jeopardy.

Due to the already heavy communication load placed on operators and the need to communicate with a wide group of people, the use of real-time, text-based communication (Chat) has proliferated in C2 centers (Eovito, 2006). Warfighters reported that Chat is an essential tool to accomplish their mission (Eovito, 2006). They stated that Chat is invaluable because it is a fast and efficient way to communicate to multiple people without having to transmit on multiple radio channels or repeat information. Another advantage of Chat is that it creates a log of all transmissions, which operators may reference for clarification while at the same time eliminating the need to take notes, thus reducing the chance for errors in the dictation of messages. The fast and wide spread dispersion of information builds greater

situation awareness amongst team members since they are constantly updated on the current state of the mission. Although action by the operators may not be needed immediately, the improved situation awareness of the current mission allows the operator to act faster since they now have a more complete picture of the task.

Enhancement of shared situation awareness and the increased speed of C2operations are motivations behind the creation of a dense network of sensor information and weapon platforms labeled as network-centric warfare (Alberts, Garstka, & Stein, 1999). The idea is that real-time distribution of information to the operator fosters self -synchronization and essentially improves performance in all areas of operation. Chat has proven to be a useful tool for network-centric warfare, however there is a lack of interoperability between communication systems forcing operators to divide their attention between multiple displays. Studies have shown that dividing attention between multiple displays, although providing rich, real-time data, can cause sub-optimal performance on primary task performance (Cummings & Guerlain, 2004). To help improve communication performance for network-centric warfare, researchers in the Air Force Research Laboratory's Battlespace Acoustics Branch have developed a network-centric communication management suite. This system, called Multi-Modal Communication (MMC), combines advanced communication technologies into an integrated communication system. As seen in figure 1, MMC fosters collaborative decision-making by integrating Radio and Chat communication data, from distributed operators, in an intuitive display. The following section describes the specific tools that make up MMC.
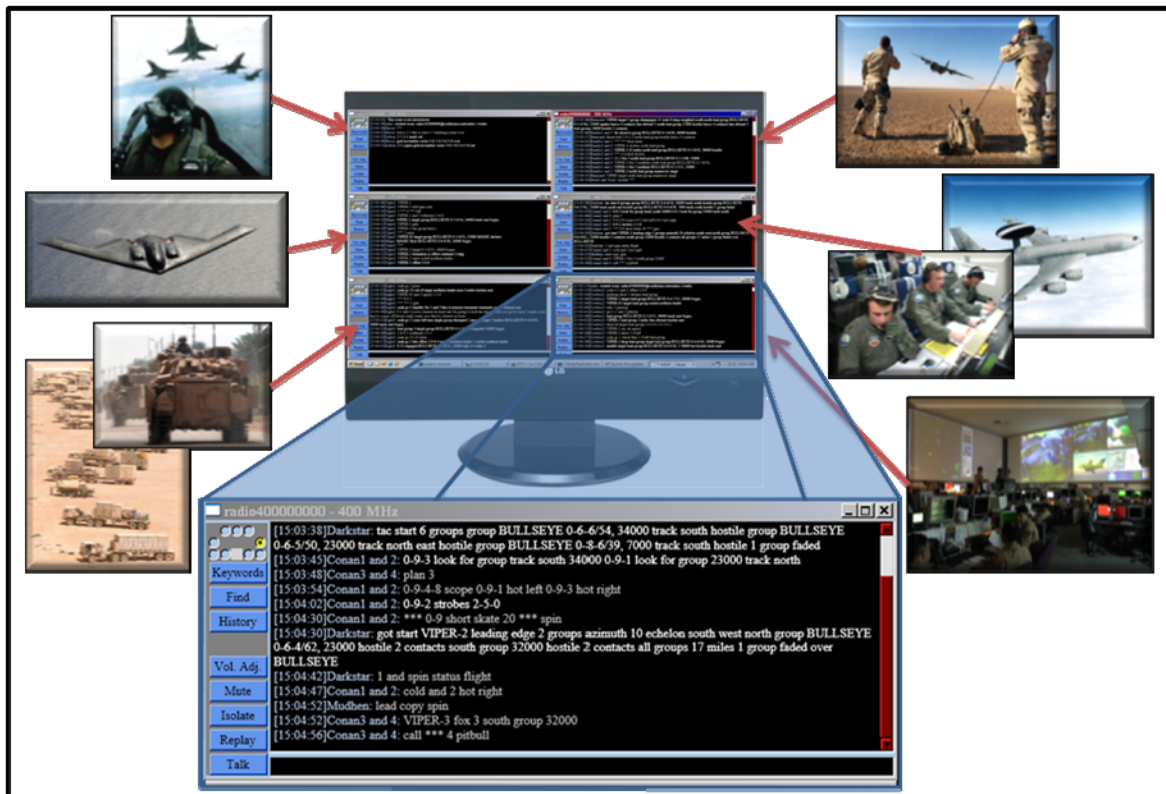

Figure 1. Depiction of the Multi-Modal Communication display managing communication data from multiple distributed operators.

MMC Features

MMC allows operators to manage communication from voice as well as text-based systems in a single intuitive, dynamic display. A number of tools are built into MMC, which allow operators to easily find and retrieve past information. In addition to aiding in the retrieval of information, speech intelligibility over the Radio channels is increased by spatially separating each of the radio channels to virtual locations around the operator via their headphones. This Radio data is also captured and recorded so that it is available to be replayed. MMC also transcribes the Radio data and displays it as text thus creating a persistent depiction of verbal communication. The combination of these features provides operators with the tools necessary to monitor multiple communication channels for critical information and make accurate decisions quickly. The following sections provide a more detailed view into the functionality of MMC.

Spatial Audio

Several studies have shown that spatializing talkers into virtual locations increases intelligibility by 30-40% for headphone-based multi-talker communication (Bolia & Nelson, 2003; Brungart & Simpson, 2005; Ericson, Brungart, & Simpson, 2004; Nelson, Bolia, Ericson, & McKinley, 1998). A large part of this outcome is due to the "cocktail party effect" (see Cherry, 1953), which refers to the fact that at a cocktail party a person is inundated with multiple conversations yet can focus on an individual one as well as shift their attention to another at will. However if a person were to listen to the same conversations over monaural headphones the effect would be lost thus the spatial separation of simultaneous talkers largely contributes to this effect. Digital filters, called head-related transfer functions (HRTF), take advantage of binaural cues that normally occur when competing talkers are spatially separated, and displays the sounds over headphones in a similar manner. The use of the HRTFs has lead to the development of auditory displays, which produce virtual sounds that are as localizable as sounds produced in the free field. Such auditory displays have been shown to increase speaker intelligibility, and reduce listeners' perceived mental workload (Bolia, 2003; Crispien & Ehrenberg, 1995; Martin, McAnally, & Senova, 2001; Nelson, Bolia, Ericson, & McKinley, 1999). A study by Brungart and Simpson (2005) systematically evaluated different configurations of HRTF filters and found an optimal set that maximizes the number of talkers while preserving intelligibility. MMC uses these filters to allow operators to place the radio channels in one of nine virtual, spatial locations and to change that location in real-time during the mission. The flexibility of this configuration allows the operator to organize and more efficiently monitor multiple radio channels. Figure 2, depicts the virtual spatial locations from which an operator could hear verbal communication.

Figure 2. Representation of the virtual sound source based on selection in the Spatial Audio control. A user selects the location they wish to hear the communication to virtually originate from and the system selects the corresponding HRFT to filter the audio stream.

Capture and display of messages

As previously stated, a major issue with verbal communication is that it is perishable, thus operators only have one chance to hear a message before it is gone. Communication effectiveness, or the likelihood that speech communication will be accurately received, can be degraded by numerous factors including ambient noise in the environment, loss of attentional focus by the listener, and masking by competing speech signals (Bolia, 2003). In order to maximize communication effectiveness, the MMC employs real-time speech-to-text transcription for all verbal communication and display the transcribed speech as text in a Chat-like window thus preserving the message for later recall. This consequently reduces the need for note taking since the message displayed at text, as seen in Figure 3. In addition to transcribing the message, MMC also stores the original audio transmission and links it to the text. Operators simply need to select a line or lines of text to relisten to the message (highlighted phrase in Figure 3 is the selected line to be replayed). MMC also has a REPLAY function, which replays the last 15 seconds of communication. An operator would use the REPLAY function in cases where they hear something important or are unsure of what they actually heard, and want to hear it immediately repeated without having to look for it in the display window. The ISOLATE function was added to the MMC suite to allow operators to mute all other channels so they can give their undivided attention to one specific channel. The ability to read verbal messages or hear them again reduces the need to request a message to be repeated, thus saving time and communication bandwidth but more importantly provides the operator with easy access to accurate information.
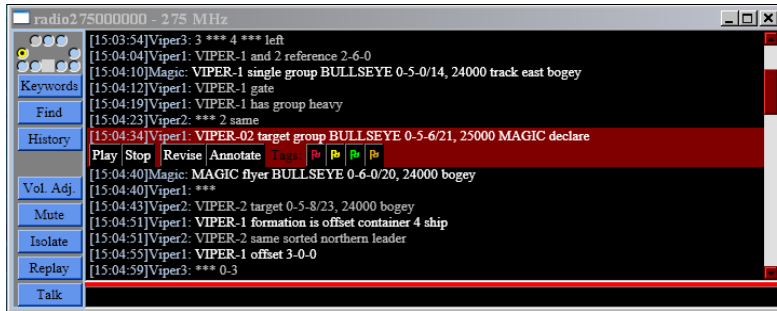
Figure 3. Display of text transcriptions of verbal messages. The highlighted phrase is selected for the audio message to be replayed.

In addition to verbal radio messages, MMC also accepts and transmits Chat messages. Similar to most multi-user Chat programs, MMC follows a client-server framework where a user sends a Chat message via the client to a server. The server then distributes the Chat message to the connected clients and stores both the Chat data and transcribed voice data in time order, thus maintaining a complete log of all communication during a mission. Operators who join the communication channels after the start of the mission, who have been disconnected, or who take over a shift now have access to all verbal and written messages, thus allowing them to gain situation awareness by reviewing the log of past messages.

This log of communication affords operators full access to past information in that they can retrieve messages to confirm or clarify ambiguous data. MMC also allows users to flag lines of text for later review. As seen in the top left of Figure 4, the red FLAG is selected thus changing the text to red so that information is more salient when having to perform a search. This window also depicts the ANNOTATE function which allows users to add comment lines of communication. Annotations are stored with the original message and serves as a notepad for the operator to use during the mission or for after action review.

The FIND function in MMC allows operators to perform searches within a specific channel or over the entire server. In searching for a keyword within a channel, operators can go through each incidence of the word or pull up another window that compiles every phrase that contains the keyword, as seen in the top right and bottom left windows of Figure 4, respectively. The FIND function also allows operators to go directly to items they flagged. In addition to these search functions, operators can also use the HISTORY function which allows them to go back to a specific time, whether that is absolute or relative time, as seen in the bottom right window in Figure 4.
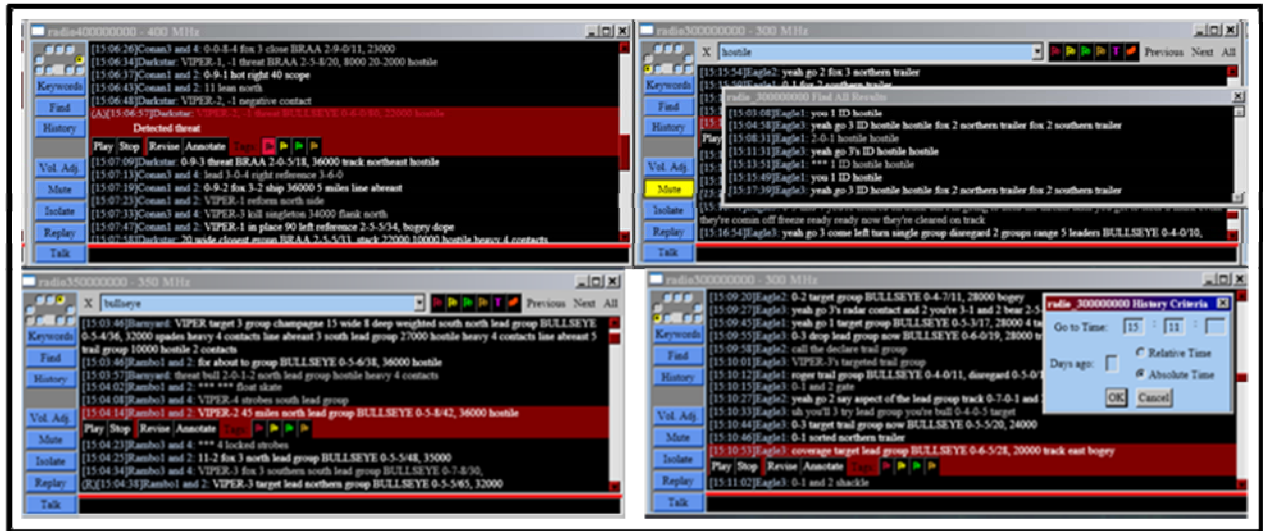
Figure 4. Four MMC display windows depicting the FIND functions. Window 1, top left, shows the FLAGGING and ANNOTATE feature. Window 2, bottom left, shows FIND feature. Window 3 top right, contains the FIND ALL feature which compiles all of the keywords into a pop up window. Window 4, bottom right, demonstrates the HISTORY function.

Edit messages

Although automatic speech recognition has greatly improved, it is still not completely accurate. For situations confronting C2 operators, where the passing of information is of the utmost importance, less than perfect transcriptions are unacceptable. MMC allows each message to be edited to reflect the correct information, as seen in Figure 5. The originator or other observers monitoring the channel can correct the mis-transcribed speech. A system can be designed so there are operators dedicated to the review and correction of errors in the transcriptions, thus serving as an editor to the transcriber to maintain accurate transcriptions. The right window in Figure 5 shows a list of all past corrections and the original audio file is still attached to the corrected message, allowing the user to verify the message.
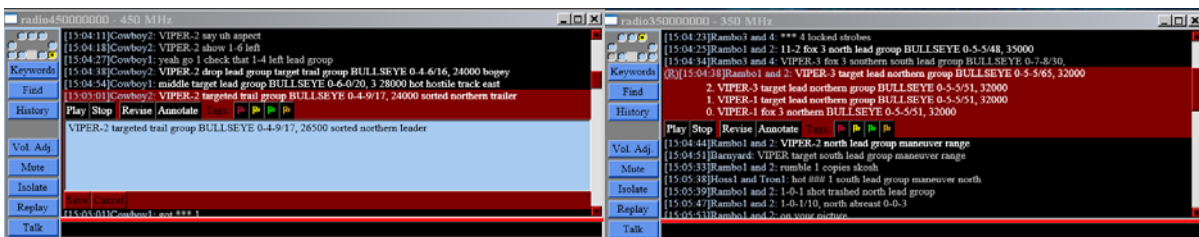

Figure 5. MMC windows displaying the edit feature.

Keyword

Operators are inundated with communication, which they must constantly monitor for information that is pertinent to them (Bolia, 2003). Having to monitor multiple streams of communication, as well as other types of displays, creates situations in which it is very easy to miss critical information. MMC allows operators to set up alerts for keywords that are instrumental to their mission. As seen in Figure 6,

the selected KEYWODS that are configured by the operator are highlighted as they appear in the window, thus alerting the operator to that information.
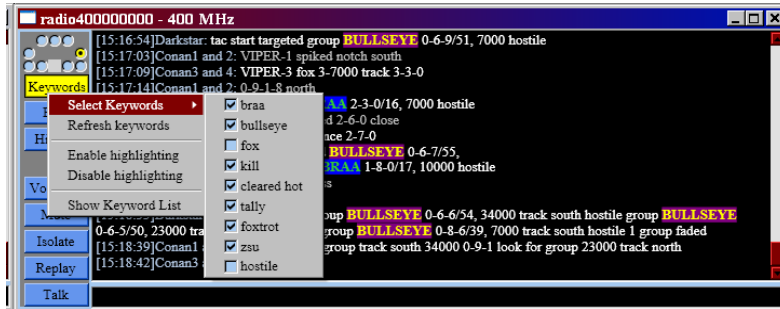

Figure 6. MMC window with keyword selection.

Evaluation of the Multi-Modal Communication Tool

This study is an initial evaluation of the MMC tool. The goal was to assess the potential utility of the MMC suite as a communication-management tool, as compared to standard Radio communication, Radio communication with 3D spatial audio, and Chat. It was hypothesized that the combination of communication tools would aid operators in their ability to quickly and accurately detect and reply to critical messages. In addition to manipulating communication devices, the amount of information that participants had to monitor was also varied. The "Short" condition required that operators monitor speech communication for critical signal comprising a call sign and the word "hostile." And, they were only required to repeat the call sign and the word "hostile." Conversely, in the "Long" condition operators were required to respond with the all of the target's location information, thus requiring more of a working memory load on the operator. The second hypothesis is that the combination of communication tools would aid operators in their response to the Long response phrases.

 Methods

Participants

Ten paid participants, five male and five female, ranging in age from 18 – 26 years old served as operators in this task. All operators had normal hearing, as well as normal or corrected-to-normal vision.

Design

A within-subject design study was employed, with two signal difficulties (Short and Long) combined factorially with four levels of communication interfaces (Radio, 3D-Audio, Chat, and MMC). The Radio condition consisted of labeled channels with a push-to-talk button. Audio from all six channels were presented monaurally. The 3D-Audio condition spatially separated each audio channel, so that each one appeared to emanate from a different spatial location around the operator. The Chat condition sound was not transmitted, but instead text transcriptions of the signals were displayed for the six channels. Finally, in the MMC condition, the operator was able to hear each of the channels localized in space and

see the speech-to-text transcription of all audio signals. In addition, several other interface tools were made available to the operator, including the mute, isolate, playback, and search function.

Apparatus

In all experimental conditions, operators participated in a 10-min communication-monitoring task. The participants simultaneously monitored six Radio/Chat channels for the occurrence of critical phrases. Each channel had a specific call sign assigned to it. The critical and neutral phrases for each call sign were recorded by six different voices (four male and two female), with one voice assigned for each call sign so that every channel had a distinct call sign and voice. Critical signals were defined by the presence of the word "hostile." In the Short condition, the critical phrase contained the channel's call sign followed by the phrase "ID hostile" (e.g., "Viper one ID hostile"). The critical phrase for the Long condition was the call sign for the channel followed by the word "hostile" and location information (i.e. "Viper one hostile north lead group 25 miles"). The neutral phrases were modified phrases from the Air Force Tactics, Techniques and Procedures communication brevity document that did not contain the word "hostile."

Signals for each of the six channels were updated independently from each other. Neutral signals were generated randomly over a range of 8 – 30 seconds, with the restriction that there were four neutral signals per channel per minute (240 overall neutral signals). The message length for the neutral signals ranged from 1 – 8 seconds long (*M* = 3.85 sec). Critical signals varied at random over a range of 15 – 120 seconds, with the restrictions that critical signals were generated on average once per channel per minute (60 overall critical signals; signal probability = .25). The message lengths for the short critical signals were on average two seconds long while the long signals ranged from three to five seconds with an average of 4.3 seconds. In all conditions, observers responded to a critical signal by pressing the "push-to-talk" button on the correct channel and repeating the message back. Responses occurring within 50 seconds of a critical signal were recorded and all responses beyond were considered a missed signal.

Procedure

All participants took part in the computer-based training, which explained that their task was to monitor communication channels for the occurrence of a hostile entity that required a response. The training, which was comprised of two separate hour and a half sessions, provided participants experience with the different communication interfaces and the task. The eight experimental conditions were randomized per participant. Data was collected in a series of sessions, with each session usually lasting about 20 minutes. Each participant typically participated in one session a day, until the completion of all eight conditions. Participants were stationed at a computer, where they experienced the presentation of neutral and critical signals through the communication interface. Headphones were used to present the auditory signals and the participant's responses were recorded through the attached microphone.

Results

Correct Detections. Mean percentage of correct detections for all eight conditions are presented in Figure 7. The data for Figure 7 were tested for statistical significance by means of a 2 (difficulty) × 4 (comm. device) within-subjects analysis of variance (ANOVA). Significant main effect for difficulty was found, F (1, 9) = 26.44, $p < .05$, in that detection, as measures by the number of responses to critical phrases, were greater in the Short ($M = 81.3$) than in the Long condition ($M = 67.9$). A significant main effect was also found for communication device, F (1.56, 14.02) = 26.44, $p < .05$, where detection performance was greater for MMC ($M = 84.3$) and Chat ($M = 83.6$) which were not different from each other but where greater than 3D Audio ($M = 69.8$) and Radio ($M = 60.7$). 3D Audio had significantly more detections than Radio. A significant interaction between difficulty and comm device is illustrated in Figure 7, F (1.69, 15.23) = 5.97, $p < .05$. Post hoc tests were performed to explore the significant interaction. Detections were greater in the Short than Long condition for the Radio and 3D Audio conditions however, there was no difference in difficulty for the Chat and MMC conditions. In this and all subsequent ANOVAs, Box's epsilon was used to correct for violations of the sphericity assumption (Maxwell & Delaney, 2004).
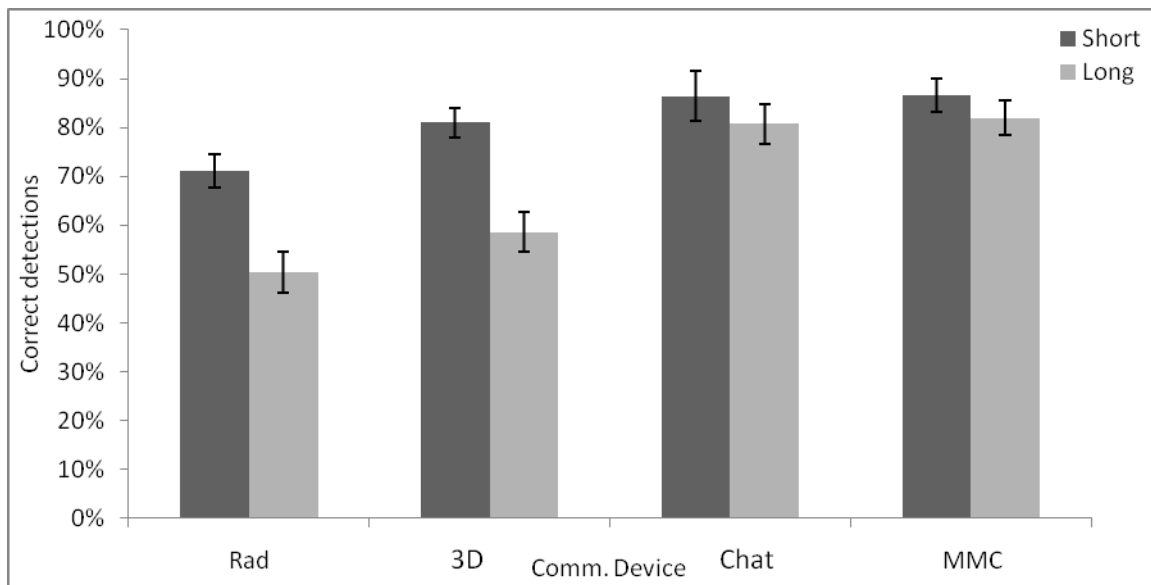

Figure 7. Mean percentage of correct detection as a function of difficulty and communication device.

Accuracy. Mean response accuracy for all eight conditions are presented in Figure 8. Accuracy scores were calculated from the number of words in the response that were identically matched to the words in the critical phrase.
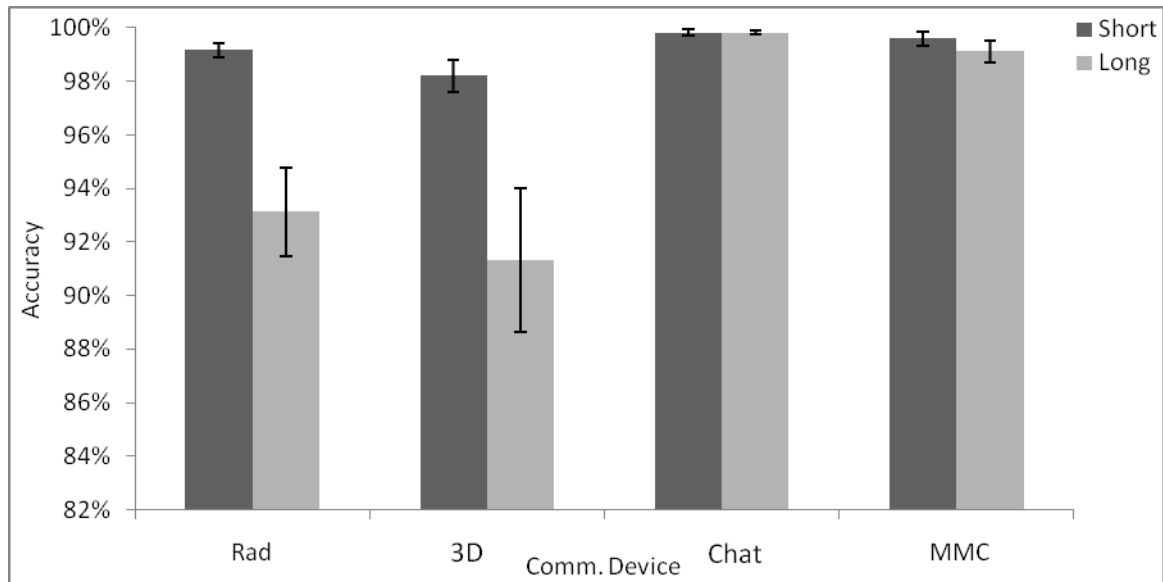
Figure 8. Mean accuracy scores as a function of difficulty and communication device.

A 2 (difficulty) × 4 (comm. device) within subjects ANOVA revealed a significant main effect for difficulty, F (1, 9) = 11.17, p < 05, was found for response accuracy in that the Short condition (M = 99.2) was greater than in the Long condition (M = 95.8). A significant main effect was also found for comm. device, F (1.14, 10.24) = 9.96, p < 05, where accuracy was greater for MMC (M = 99.4) and Text (M = 99.8) which were not different from each other but there greater then 3D Audio (M = 94.8) and Radio (M = 96.1) which too were not different from each other. A significant interaction between difficulty and comm. device in illustrated in Figure 8, F (1.18, 10.62) = 8.07, p < .05. Similar to detections, post hoc tests revealed that accuracy was greater for the Short than Long phrases for the Radio and 3D-Audio conditions however there was no difference in difficulty for the Chat and MMC conditions.

Reaction Time. Response times of the correct detections for all eight conditions are displayed in Figure 9. Again, a 2 (difficulty) × 4 (comm. device) within subjects ANOVA on the response time data found a significant main effect for difficulty, F (1, 9) = 65.93, p < 05, in that the Short condition (M = 4.39) was faster than in the Long condition (M = 8.82). A significant main effect was also found for comm. device, F (2.02, 18.24) = 29.52, p < 05, where response time was fastest for Radio (M = 3.54) and 3D Audio (M = 3.35) which were not different from each other but were faster than MMC (M = 6.98) which in turn was faster than Chat (M = 12.54). The interaction between difficulty and comm. device was not found to be statistically significant, p >.05.
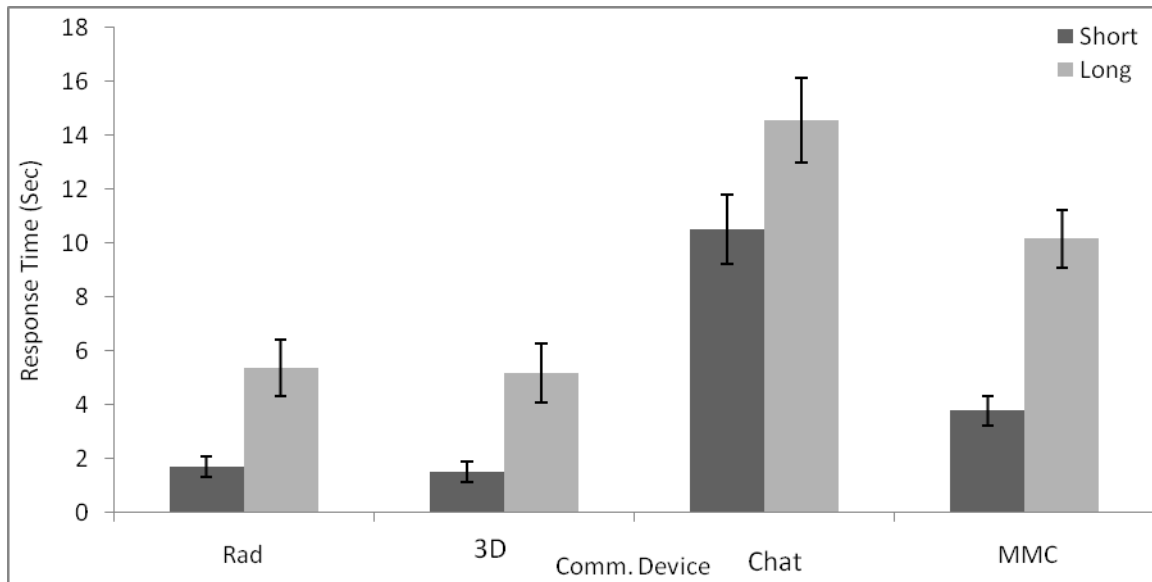
Figure 9. Mean response times in seconds as a function of difficulty and communication device.

Discussion

This was the first study to evaluate the MMC tool to aid C2 operators in managing communication data. MMC is a network-centric communication management suite that integrates mature communication tools into a single intuitive display. The integration of tools into such a display aids operators in their ability to monitor multiple communication channels to extract critical information. The study was designed to evaluate different communication tools and the extent these tools aid operators in monitoring and replying to critical messages. Operator's ability to quickly and accurately reply to detected signals was used to evaluate their performance with each of the communication devices.

The two most common types of communication systems in current use are Radio and Chat. These results shown here clearly indicate that messages were easier to detect and could be read back more accurately with a chat-based communication system, but that listeners were able to respond more rapidly to critical information with a radio display. Thus, in choosing to present information through either a Chat or a Radio display, there is a clear speed versus accuracy tradeoff between these two alternatives. The proposed MMC system offers a compromise between these two extremes - it provides a level of accuracy comparable to that achieved with a Chat display, but with a much faster response time (although it was not quite as fast as a Radio display). The most likely explanation for this result is that listeners in the MMC condition would respond immediately in cases where they were sure they heard the information correctly, but that they would wait until the verbal transcription of the text occurred in cases when they were uncertain about the contents of the critical message. Another contributing factor may be that the verbal message drew their attention to the correct chat window when new information was arriving, which allowed them to respond more quickly than they were able to in the Chat-only mode where they had to constantly scan all the chat windows for new information.

12

Another important finding was that there were no differences in correct detections and response accuracy for difficulty in the Chat and MMC conditions whereas there were differences in the Radio and 3D-Audio conditions. This suggests that the availability to read messages aids in the monitoring of communication, independent of message complexity. Situations such as the ones confronting C2 operators where the transmission of complex data is of utmost importance, the availability of persistence transcriptions of Radio communication and the ability to easily retrieve pertinent information could be a matter of mission success or failure.

Conclusion

Currently Radio is still the standard form of communication and comparing operators' ability using radio as opposed to the MMC tool, it is clear the added functionality of MMC greatly improves the operators' performance. Although operators were faster in their responses with the Radio as compared to MMC, participants using the MMC tools detected 24% more messages and were more accurate in their responses then with the use of the Radio. However, the comparison between Chat and MMC is not as clear. There was not a significant difference between Chat and MMC for messages detected and response accuracy which suggests that Chat alone aids in communication monitoring. However, there was a difference when looking at the time it took operators to respond to a critical message, operators were significantly faster with MMC than Chat. This difference can be contributed to the presence of audio priming operators to the occurrence of a critical phrase. In addition, the persistent nature of Chat allowed operators to respond to messages they might have previously missed with audio messages alone.

This was the initial evaluation of the MMC tool and although it was a simple message detection task there is evidence that the integration of communication tools aids operators in their ability to monitor multiple communication channels. These results suggest MMC provides a balance between the speed of Radio listening and the increased intelligibility of 3D-audio display with the data capturing capability of Chat only displays. A survey of 26 air battle managers from the United States Air Force, United States Navy and the Royal Australian Air Force assessed the perceived usefulness of current commercial-off-the-shelf collaborative tools (Nelson, Bolia, & Vidulich, 2004). The highest ratings for most potential technologies were data capture and replay tools and Chat and messaging, tools which are key components integrated into MMC. Based on C2 operators' wants and experimental data, the MMC tool can be a useful tool for network-centric communication management.

References

Alberts, D. S., Garstka, J. J., & Stein, F. P. (1999). *Network centric warfare: Developing and leveraging information superiority.* Washington, DC: CCRP Publication Series.

Boila, R. S. (2003). Effects of spatial intercoms and active noise reduction headsets on speech intelligibility in an AWACS environment. *Proceedings of the Human Factors and Ergonomics Society, 47,* 100-103.

Boila, R. S., & Nelson, W. T. (2003). Spatial audio displays for target acquisition and speech communications. In L. J. Hettinger & M. W. Haas (Eds.) *Virtual and Adaptive Environments:*

Applications, Implications, and Human Performance Issues. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Brungart, D. S., & Simpson, B. D. (2005). Optimizing the spatial configuration of a seven-talker speech display. *ACM Transactions on Applied Perception, 4,* 430-436.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25,* 975-979.

Crispien, K. & Ehrenberg, T. (1995). Evaluation of the "cocktail party effect" for multiple speech stimuli within a spatial audio display. *Journal of the Audio Engineering Society, 43,* 932-940.

Cummings, M. L., & Guerlain, S. (2004). Using a chat interface as an embedded secondary tasking tool. In D.A. Vincenzi, M. Moulous, & P.A. Hancock. *Human Performance, Situation Awareness and Automation: Current research and trends.* (pp.240-244). Mahwah, NJ: Lawrence Erlbaum.

Eovito, B. A. (2006). The impact of synchronous text-based chat on military command and control. *Proceedings of the International Command and Control Research and Technology Symposium, 11*.

Erickson, M. A., Brungart, D. S., & Simpson, B. D. (2004). Factors that influence intelligibility in multitalker speech displays. *The International Journal of Applied Psychology, 14(3),* 313-334.

Martin, R. M., McAnally, K. I., & Senova, M. A. (2001). Free field equivalent localization of virtual audio. *Journal of the Audio Engineering Society, 49,* 14-22.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data (2nd ed.).* New Jersey: Erlbaum.

Nelson, W. T., Boila, R. S., Ericson, M. A., & McKinley, R. L. (1998). Monitoring the simultaneous presentation of spatialized speech signals in a virtual acoustic environment. *Proceedings of the 1998 IMAGE Conference,* 159-166.

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