Automated Feedback and Situation Awareness in Net-Centric C3 Systems at Varying Difficulty Levels

John S. Barnett & Larry L. Meliza
U.S. Army Research Institute
Orlando, FL
john.barnett1@us.army.mil, larry.meliza@us.army.mil

Jennifer M. Ross
University of Central Florida
Orlando, FL
jennifer.marie.ross@us.army.mil

ABSTRACT

Net-centric command, control, and communications (C3) systems provide vital information to military commanders and increase their situation awareness (SA) of the battlespace. However, the amount of information presented and the dynamic nature of that information often makes it difficult to focus attention appropriately. Immediate feedback, in the form of automated alerts, can help direct the user's attention to important events and can be a valuable tool both for training and during military operations. Research in other areas (aviation) has shown under certain conditions automated alerts improve SA, but under other conditions alerts can interfere with SA. Theory suggests workload may be a factor. The present research sought to determine how automated feedback and workload affects SA for users of net-centric C3 Systems. In the experiment, participants viewed a laptop computer showing a simulated C3 display running a typical combat scenario, and were asked to look for and remember key events. A software application called SHIELD (System to Help Implement and Empower Leader Decisions) provided automated alerts of potentially hazardous events, such as violating unit boundaries or approaching minefields. Each participant completed two trials, one in the experimental condition and one in the control condition, with the order assigned randomly. In the experimental condition, subjects were alerted to certain situations. In the control condition, no alerts were provided. The participant's SA and workload were measured periodically during each trial. SA and workload scores were compared across the experimental and control conditions. Results showed that immediate feedback increased SA for participants with no military experience, but not for Soldiers.

Dr. John S. Barnett is a research psychologist with the U.S. Army Research Institute for the Behavioral and Social Sciences where he conducts research in training and evaluation. He holds a Ph.D. in Applied Experimental and Human Factors Psychology from the University of Central Florida. He is a former U.S. Air Force officer with a background in aviation, operations planning, and command and control.

Dr. Larry Meliza received his doctorate in psychology from the University of Arizona. He has over twenty-five years of experience as a research psychologist, working in the areas of exercise management and feedback sessions for U.S. Army collective training. Much of this work has involved automated after action review systems. His most recent work has focused on providing units with feedback regarding their application of networked command and control systems.

Jennifer M. Ross is a fourth year graduate student at the University of Central Florida in the Applied Experimental Human Factors Ph.D. program. Jennifer earned her M.A. in Modeling and Simulation from the University of Central Florida.

Key words: Automation, Feedback, Net-Centric, Situation Awareness, Training, Workload

Automated Feedback and Situation Awareness in Net-Centric C3 Systems at Varying Difficulty Levels

John S. Barnett & Larry L. Meliza
U.S. Army Research Institute
Orlando, FL
john.barnett1@us.army.mil, larry.meliza@us.army.mil

Jennifer M. Ross
University of Central Florida
Orlando, FL
jennifer.marie.ross@us.army.mil

Sharing information over a computer network is becoming a common means of communication in many fields. The U.S. Army has taken advantage of this technology to improve command, control, and communication (C3) for its forces. Field commanders can now watch events unfold on a computer screen.

The use of networked computers to enhance C3 has opened up new possibilities. Data can be analyzed by computer and results displayed to leaders, providing them with better information on which to base decisions. Leaders can be provided with immediate feedback about their performance. One promising technology is for computer systems to monitor the data stream and provide alerts when critical events occur to ensure they are not missed by the operator.

As the U.S. Army gains more experience with networked digital C3 systems, more features are likely to be added to the systems. Features such as the automated alerting system described above can help direct the user's attention to important events and increase situation awareness (SA). However, experience with automated alerting systems in other areas has shown that automation can sometimes become intrusive and capture the user's attention at inappropriate times.

This raises the question of whether immediate feedback or automated alerting systems enhance SA or interfere with SA. Research suggests that it depends on the situation and environmental variables (Billings, 1997). Under certain conditions alerts improve SA while under other conditions they interfere with SA. If immediate feedback and alerts are built into future networked C3 systems, it would be important to know how alerts affect SA, and under what conditions.

To answer these questions, we conducted an experiment using a simulation of a networked C3 system with an automated alerting application. This paper presents the experiment and our findings. The next section discusses some background about networked C3 and concerns about how automated alerts may affect SA, as well as some related concerns. Following that, we will discuss the method used for the experiment, followed by the results and a discussion of results.

BACKGROUND

Digital C3

The U.S. Army employs net-centric C3 systems in a process referred to as digitization. These systems employ computer automation to help leaders and Soldiers perform many of the C3 functions previously accomplished manually, such as planning missions, distributing orders and reports, and creating and distributing map overlays containing battlefield graphics such as obstacle areas, unit boundaries, and phase lines.

Information on the tactical situation can be distributed over the network from upper-echelon command centers down to the lowest-level combat formations, and vice-versa. Digitization not only increases combat capabilities, but also improves safety by reducing the chances of fratricide or "blue on blue" incidents. In addition, combat units who use digital systems are expected to maintain better SA and to plan and execute operations more quickly than non-digital units (Barnett, Meliza, & McCluskey, 2001).

Thus, digital C3 serves as a decision-support system for combat commanders. It helps them visualize the battle space and presents needed information in a format that fosters the commander's SA. Digitization also provides analytical tools, such as terrain analysis tools and automated warnings that can further enhance SA.

There are a number of different digital systems. Many of the systems are specific to certain Battlefield Functional Areas (BFAs) such as Intelligence, Maneuver, Field Artillery, Air Defense, and Combat Service Support. These systems were designed to fulfill C3 functions related to the BFAs and are typically located in tactical operations centers (TOC) at higher echelons. Although these systems were originally designed to operate within the BFA, they are able to share most information with other BFA systems on the network. Operators of these systems spend part of their time observing SA displays and part of their time using analytic tools and preparing products.

The Force XXI Battle Command, Brigade and Below (FBCB2) (see figure 1) is a digital C3 system designed for lower-echelon maneuver units and is typically located in vehicles such as tanks and infantry fighting vehicles. FBCB2 allows lower echelon units to plan missions and routes, develop battlefield graphics such as obstacle overlays, and share the information over the network.

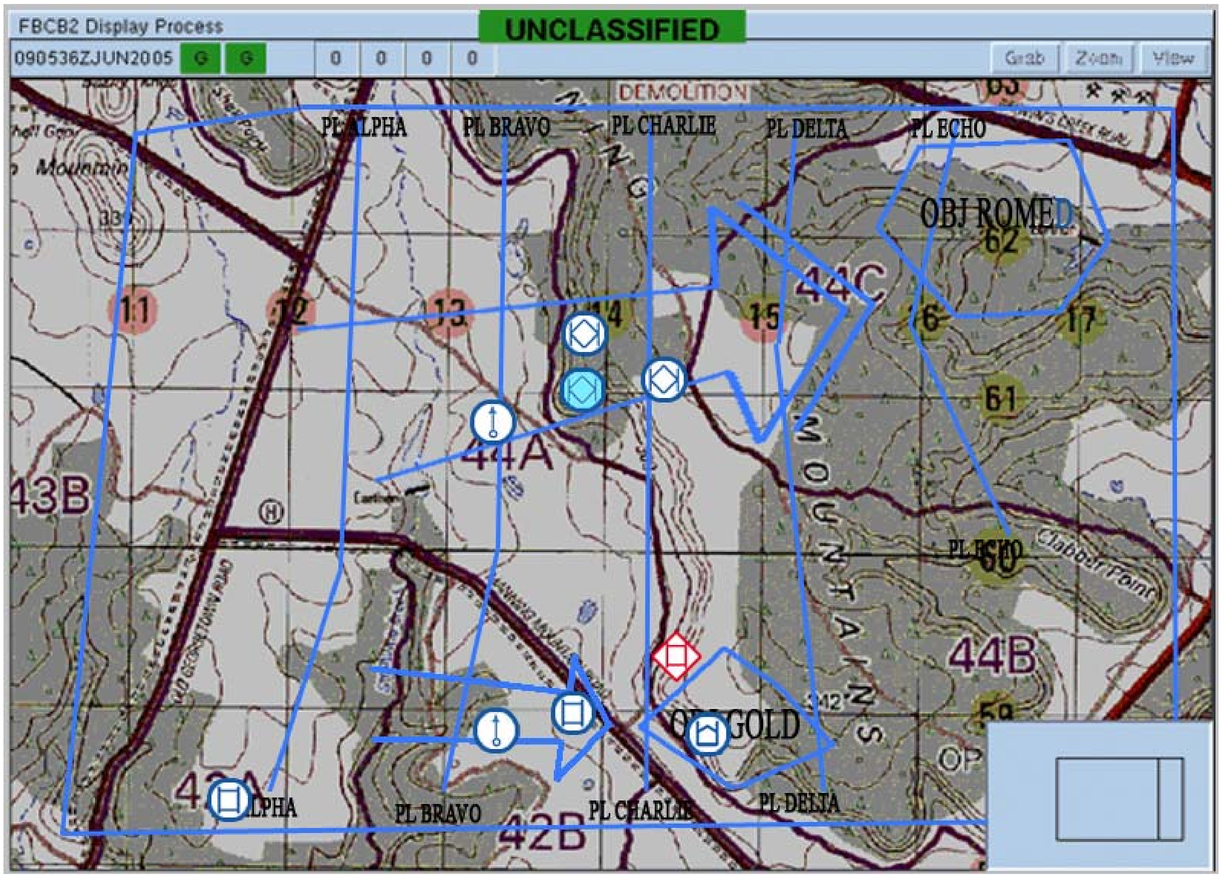

Figure 1. Example of an FBCB2 Display

FBCB2 presents a dynamic view of the battlespace. Once the mission is executed, leaders and Soldiers can follow the progress of the mission on the FBCB2 display. Vehicle-mounted FBCB2s can automatically update the vehicle's position using a Global Positioning System (GPS) position. The vehicle's position is periodically updated and transmitted to the network, so that the FBCB2 display shows the vehicle's own position as well as the positions of other vehicles in the unit. This allows leaders and Soldiers to develop sound awareness of the friendly situation. In addition, when enemy units are reported on the network through SPOT reports, their positions are displayed as well. In the field, vehicle crews are likely to spend little time viewing FBCB2 SA displays, and most of their time looking for threat situations.

Automation's Affect on Situation Awareness

Research suggests that including automation into a system introduces potentially detrimental consequences (Bowers, Deaton, Oser, Prince, & Kolb, 1995), including complacency (Morgan, Herschler, Weiner & Salas, 1993), increased monitoring requirements (Kantowitz & Campbell, 1996), and a loss of SA (Bowers, Oser, Salas & Cannon-Bowers, 1996).

A number of different factors may contribute to a reduction in SA. Bowers, Oser, Salas & Cannon-Bowers (1996) suggest that relying too much on the automation leads to complacency, which in turn leads to reduced SA. Automation may reduce workload during low-workload periods and increase workload during high workload periods (Parasuraman, Mouloua, Molloy & Hilburn, 1996) .

Therefore, while automated alerts may be helpful during low workload periods and help increase SA, during high stress, high workload periods, alerts may become intrusive and responding to them may not only increase workload, but shift the operator's attention from more important tasks, thus reducing SA.

Change Blindness

Another consideration for introducing automated alerting into digital C3 systems is that sometimes operators who view visual displays will fail to detect changes that occur on those displays, a phenomenon called change blindness. Change blindness tends to occur concurrent with various types of visual transients such as icon movement, screen flashes, or eye blinks. In addition, operators can fail to detect changes if they are performing other tasks. Durlach and Chen (2003) found individuals tended to detect changes in icon appearance, disappearance, and color changes, but had more difficulty detecting changes in icon type and movement, particularly if the icon was in the periphery of the screen and movement was small. The concept of change blindness relates to automated alerts in two ways. First, alerts may bring critical events to the attention of system operators at points in time when SA displays are not being closely monitored. Second, alerts may hinder the operator by capturing the operator's attention at inopportune times, and thus causing the operator to miss a change on the display they would otherwise have noticed. Either case would have an affect on the operator's SA.

System to Help Implement and Empower Leader Decisions

The System to Help Implement and Empower Leader Decisions (SHIELD) was developed under a Phase II Small Business Innovation Research contract (Aiken, Green, Arntz, & Meliza, 2005). SHIELD has numerous features designed to reduce intrusiveness and support the conduct of after action reviews (AARs). SHIELD was designed to run as a stand alone system or as an application on any network system. It has been demonstrated running as an application on FBCB2 and the Command and Control PC (C2PC). This feature allows the work of monitoring alerts to be distributed among nodes within a network (e.g., SHIELD within the TOC may provide an alert when any platform within the unit approaches a minefield and/or a SHIELD running on vehicle may alert the vehicle commander when his/her vehicle approaches a minefield). SHIELD allows the user to decide whether alerts are triggered from a unit or vehicle perspective. SHIELD provides information in several different formats, including textual and graphic displays, as illustrated in Figure 2. SHIELD allows the user to temporarily dismiss an alert or have the alert go away for the rest of a mission. SHIELD maintains an AAR log file that enables the user to call up alerts, data on user responses to alerts, and other information relevant to AARs. Users can dismiss alerts during high workload periods and then call up the AAR log file during subsequent lower workload periods to see if any of the alerts are still relevant.

Current Research

The current research seeks to answer questions about how immediate feedback/automated alerting systems affect SA under varying workload conditions. Theoretically, under conditions of low workload alerts should direct user's attention to important events and thus improve SA. Therefore, our first hypothesis is that under low workload SA will be greater with alerts enabled than with alerts disabled. On the other hand, under high workload conditions alerts may divert attention from more important tasks and consequently may interfere with maintaining SA. Thus, our second hypothesis is that under high workload conditions, SA will be lower with alerts enabled than with it disabled.

METHOD

Participants

Participants were U.S. Army Soldiers and university students. Six Soldiers and twelve students participated. For the Soldiers, mean age was 26 (SD = 8), while the mean age for students was 27 (SD = 5). The Soldiers averaged 7 (SD = 5) years of military experience, while all of the university students reported no military experience.

Apparatus

A simulation of a networked C3 system was presented on a laptop computer (Pentium M) using a 14" graphics monitor operating under 1400 by 1050 pixel resolution. The system was integrated with headphones that the participants wore during all experimental and practice trials. The C3 system simulated was the Force XXI Battle Command, Brigade and Below (FBCB2). Like FBCB2, the simulation presented a map display showing locations of friendly units, enemy units (if known) and battlefield graphics such as phase lines, unit boundaries, and obstacle belts, etc. For the purposes of this experiment three Army training scenarios were programmed, a practice scenario and

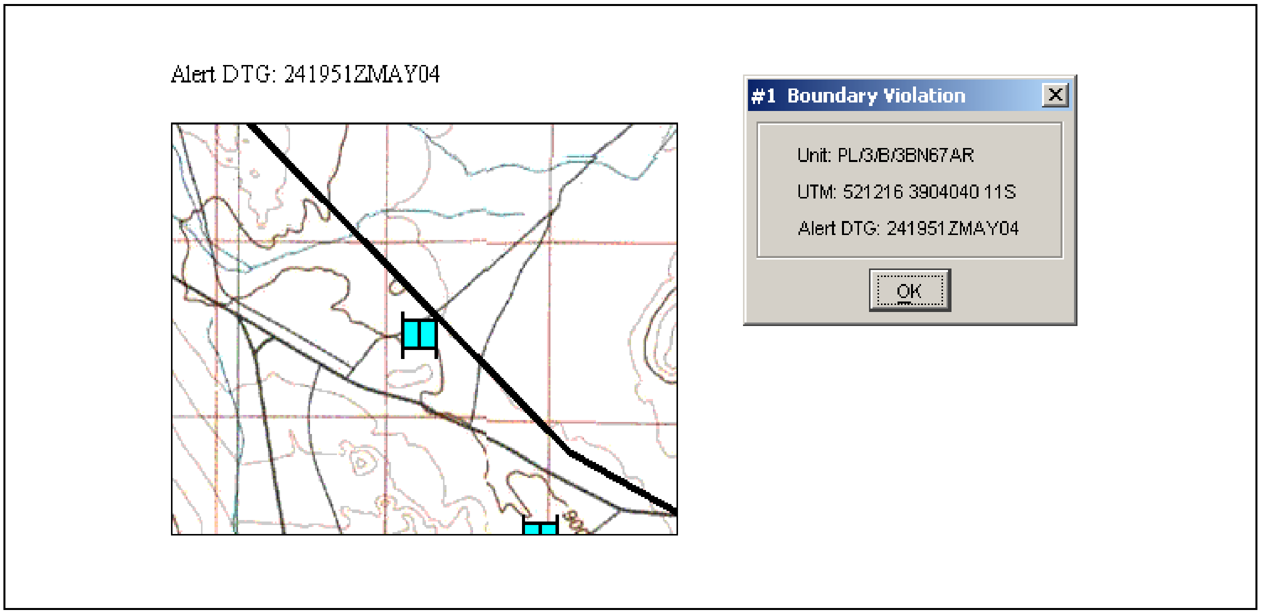

Figure 2. Examples of a SHIELD Graphic Alert (Left) and Text Alert (Right)

two full-length experimental scenarios, each based on different typographical maps and order of events.

Running concurrently with the FBCB2 task during experimental trials, was a simulation of (SHIELD. The current experiment employed SHIELD alerts triggered by five significant events. These events were unit approaching a minefield, receipt of an enemy SPOT report (i.e., report providing information about the location of enemy forces), a new friendly unit appearing on the display, a friendly unit violating battlefield boundaries, or a unit approaching a nuclear-biological-chemical (NBC) contaminated area. The SHIELD system has a number of other features which were not used for this study as they would have added confounding variables to the experiment. This experiment employed SHIELD textual alerts, but it did not include graphical alerts or other SHIELD capabilities. Further, subjects were not required to interact with the alerts in any way (i.e., subjects did not have to take actions to remove alerts from the screen), nor could they control how long the alert was displayed. The only SHIELD feature used was the visual and auditory alerts.

In each FBCB2 scenario, task difficulty, here defined as the number of significant events (i.e., minefield, spot report, etc.), was increased every five minutes. The practice condition was conducted at a low difficulty condition of 1 event roughly every 40 seconds (10 total events for a six-minute practice). In the experimental sessions the first 5 minute interval represented a low difficulty condition and had on average 1 event every thirty seconds (10 total events). The second 5 minute interval represented the moderate difficulty condition and had on average 1 event every twenty seconds (15 total events). The third 5 minute interval represented the high difficulty condition and had on average 1 event every ten seconds (30 total events). The order of significant events was randomized for each scenario with the constraint that each type of event appeared equally often as the other events during each of the five-minute blocks.

Participant's SA was measured using the Situation Awareness Global Assessment Technique (SAGAT). The SAGAT questionnaires were developed following the guidance in Endsley (2000). Example queries include recalling the approximate number of friendly units currently on the display or drawing conclusions on which objective the commanding unit is heading towards.

Subjective workload was measured using a pen and paper version of the NASA Task Load Index (TLX; Hart & Staveland, 1988). The NASA-TLX uses six dimensions to assess mental workload: mental demand, physical demand, temporal demand, effort, performance, and frustration. Each dimension is first rated by the participant on a scale from 0 to 100 with higher numbers reflecting greater workload. Next, paired comparisons are conducted, which require the participant to choose which dimension was more relevant to workload across all pairs of the six dimensions (15 total paired comparisons).

Procedure

After first completing an informed consent and demographics questionnaire, participants then completed a training session, which informed them about their task during the experiment and introduced them to the basic information they would need on FBCB2 and automated alerts. Following training participants were given a training evaluation; this evaluation ensured that they understood their task and the basic information on FBCB2 and SHIELD that they would need to complete their task. After successfully completing training participants were given a six-minute practice scenario followed by SAGAT and NASA-TLX questionnaires to familiarize them with how the experiment would proceed.

Participants then completed two experimental sessions of approximately 15 minutes each. During the sessions they were asked to view the FBCB2 display and monitor it for significant events. For one session, alerts were enabled, while for the other session alerts were disabled. The conditions (alerts enabled or disabled) and scenarios (which of the programmed FBCB2 scenarios they observed) were counterbalanced, creating four possible orders.

At pre-programmed intervals of every five minutes the FBCB2 simulation was halted and the display replaced by a blank black screen. Immediately, after the simulation was stopped SAGAT and NASA TLX questionnaires were administered to participants. After the participant completes the questionnaires, the monitoring of the FBCB2 display would be continued from the point at which it had been halted. These stops were repeated three times for each fifteen minute scenario, this number was chosen in accordance with findings by Endsley (2000) that purport no ill effects on performance due to brief pauses to access SA up to three times in a fifteen minute period. In all, there were three SAGAT and three NASA TLX measures for each FBCB2 session, for a total of six of each measure for the two sessions. Following the two sessions, participants were asked to complete an exit questionnaire which queried participants on their opinion on how alerts may impact the performance of others or themselves when using FBCB2.

RESULTS

As a manipulation check we first examined the two scenarios that were used in FBCB2, collapsed over alert condition, to check for equivalence. The two scenarios used in the FBCB2 tasks, Scenario 1 vs. Scenario 2, did not significantly differ in terms of SAGAT performance (M = 88.70, SE = 0.01 vs. M = 88.90, SE = 0.01 respectively), $t(17) = -0.19$, $p = .851$. The two scenarios moreover did not differ in overall workload levels (Scenario 1: M = 41.73, SE = 4.59 vs. Scenario 2: M = 46.82, SE = 4.27), $t(17) = -1.51$, $p = .15$.

In examining SAGAT performance, we used repeated measures ANOVA, with a 2 (Alert Condition) by 3 (Task Difficulty) design with all within-subject measures. There was a significant effect for task difficulty, $F(2, 34) = 8.04$, $p = .001$. Results indicated that the low difficulty condition had significantly better SAGAT scores (M = 93.3, SE = 1.4) then either the moderate difficulty (M = 87.8, SE = 1.7) or high difficulty (M = 85.3, SE = 1.7) sessions, which did not significantly differ from one another ($p > .05$). The effect of alerts and the interaction between alerts and task difficulty were not significant ($p > .05$ in both cases).

In regards to our first hypothesis, that under low workload SA will be greater with alerts enabled than with alerts disabled, our results using a paired comparison t-test were not significant, albeit in the right direction, $t(18) = 0.31$, $p = .186$ (see Table 1). In regards to our second hypothesis concerning the high workload condition, in which we hypothesized that alerts may divert attention from more important tasks and consequently interfere with maintaining SA, we also had nonsignificant results, $t(17) = -0.27$, $p = .793$ (although again our results were in the right direction, Alerts enabled: M = 84.94, SE = 2.53 vs. Alerts disabled: M = 85.67, SE = 0.17). The moderate difficulty session also failed to reach significant between the alerts and no alerts conditions ($p > .05$).

Table 1. SAGAT Performance as a Function of Feedback and Difficulty Condition

Feedback Condition	Difficulty Level	
	Low Difficulty	High Difficulty
Alerts Enabled	93.67 (1.61)	84.94 (2.53)
Alerts Disabled	93.00 (1.91)	85.67 (1.71)

The results suggested there may be an intervening variable which affected the analysis. On examining the data, we noticed a difference in SA scores between Soldiers and non-Soldiers. When we parsed these data into Soldier and non-Soldier and tested them, we found no significant difference between Soldier's SA scores ($t(5) = -0.44$, $p = .677$), but a significant difference between non-Soldier SA scores ($t(11) = 2.49$, $p = .03$). Those without military experience had significantly higher situation awareness with alerts enabled than without alerts (see Table 2).

Table 2. SAGAT Performance as a Function of Military Experience

Feedback Condition	Military Experience	
	Soldier	Non-Soldier
Alerts Enabled	88.47 (2.12)	89.99 (1.27)
Alerts Disabled	89.36 (2.31)	87.50 (1.64)

To examine workload we used a repeated measures ANOVA, with a 2 (alerts condition) by 3 (task difficulty) design examining NASA-TLX overall weighted workload scores with all within-subject measures. There was a significant effect for task difficulty, $F(2, 34) = 30.19$, $p < .0005$. The low difficulty condition ($M = 34.07$, $SD = 4.13$) was significantly lower workload then the moderate difficulty ($M = 43.84$, $SD = 4.08$) or high difficulty ($M = 54$, $SD = 4.84$), which were also significantly different from each other ($p < .01$ in all cases). However, the effect of alerts and the interaction between alerts and task difficulty did not reach significance ($p > .05$ in both cases). At a lower level of analysis, using paired comparison t-test between alert conditions at each one of the difficulty levels, a significant effect was found for the low difficulty level. A *t*-test revealed that there was a significant difference between the alerts enabled condition and the alerts disabled condition, $t(17) = -2.07$, $p = .05$. The results demonstrated that the alerts condition had a significantly lower workload ($M = 31.15$, $SE = 3.69$) than the no alerts condition ($M = 37.00$, $SE = 4.94$) in the low difficulty level. The moderate and high difficulty session however failed to reach significance in paired comparisons *t*-tests between the alerts and no alerts conditions for workload ($p > .05$ in all cases).

DISCUSSION

Alerts are important in bringing critical tactical situations to the attention of operator and users of networked C3 systems, however, attention must be given to the relationship between workload and the ability of alerts to enhance SA. In the current experiment, task difficulty was operationally defined in terms of the frequency of occurrence of events described by the experiment to the participants as being important to SA.

It was found that perceived workload increased as task difficulty was intensified. The benefit of lower difficulty levels was further supported by the discovery that participants demonstrated greater SA scores at the low difficulty level.

The results indicated that alerts did not impact SA and that the added load of the automated feedback also did not impact perceived workload for the group of participants as a whole. Indeed, though the perceived workload difference was not significant the mean workload for the FBCB2 tasks with alerts enabled was actually lower ($M = 41.57$; $SE = 4.01$) then the mean workload for the FBCB2 task alone ($M = 46.98$; $SE = 4.81$).

However, when the participants were grouped into Soldier and non-Soldier, differences became apparent for the non-Soldier group. The SA scores for non-Soldiers differed depending on whether alerts were enabled or not. Since Soldiers were more familiar with the task and environment simulated by FBCB2 than the non-Soldiers, alert condition may have had less of an effect on Soldier's SA, and a greater affect on non-Soldiers SA. This may suggest that expertise plays a role in how automated alerts affect SA.

If so, it suggests that including automated alerting systems similar to SHIELD in net-centric C3 systems may be more beneficial for novices vice experts. It also suggests that since automated alerting systems have no significant affect on the SA of more experienced operators, consequently it does not significantly reduce their SA.

Additional findings that may be of interest include SA and workload at the low difficulty condition. SA was significantly better and workload was significantly lower in the low difficulty condition than at moderate or high difficulty. Taken together, these findings reiterate the connection between SA and mental workload. Fracker (1989) suggested both SA and mental workload require the same cognitive resources (attention), so that increased mental workload may reduce the individual's ability to maintain SA.

It should be noted that the results presented here are only preliminary and conclusions should be approached with caution. Additional data collection is planned which may address some of the questions which remain in the present research.

REFERENCES

Aiken, D.S., Green, G.E., Arntz, S.J., & Meliza, L.L. (2005). *Real time decision alert, aid and after action review system for combat and training* (ARI Technical Report 1165). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD A 437006).

Barnett, J. S., Meliza, L. L. & McCluskey, M. R. (2001). *Defining digital proficiency measurement targets for U.S. Army units* (ARI technical report 1117). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Billings, C. E. (1997). *Aviation automation: The search for a human-centered approach*. Mahwah, NJ: Lawrence Erlbaum.

Bowers, C. A., Deaton, J., Oser, R. L., Prince, C. & Kolb, M. (1995). Impact of automation on aircrew communication and decision-making performance. *International Journal of Aviation Psychology 5* (5), 145-167.

Bowers, C. A., Oser, R. L., Salas, E. & Cannon-Bowers, J. A. (1996). Team performance in automated systems. In R. Parasuraman and M. Mouloua (Eds.) *Automation and human performance: Theory and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Durlach, P. J. & Chen, J. Y. C. (2003). Visual change detection in digital military displays. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, December 1-4, 2003

Endsley, M. R. (2000). Direct measurement of situation awareness: Validity and use of SAGAT. In M. R. Endsley and D. J. Garland (Eds.) *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates

Fracker, M. L. (1989). Attention allocation in situation awareness. *Proceedings of the Human Factors Society 33rd annual meeting*, 1396-1400.

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Mesh

Kantowitz, B. H. & Campbell, J. L. (1996). Pilot workload and flightdeck automation. In R. Parasuraman and M. Mouloua (Eds.) *Automation and human performance: Theory and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Morgan, B. B. Jr., Herschler, D. A, Weiner, E. L. & Salas, E. (1993). Implications of automation technology for aircrew coordination and performance. *Human/Technology Interaction in Complex Systems 6*, 105-136.