4. Dorneich, M.C., Mathan, S., Creaser, J.I., Whitlow, S.D., & Ververs (2005). "Enabling Improved Performance though a Closed-Loop Adaptive System Driven by Real-Time Assessment of Cognitive State", *Foundations of Augmented Cognition* (D.D. Schmorrow, Ed.). Mahwah, NJ: Lawrence Erlbaum Associates. pp. 621-630.

Enabling Improved Performance though a Closed-Loop Adaptive System Driven by Real-Time Assessment of Cognitive State

Michael C. Dorneich, Santosh Mathan, Janet I. Creaser, Stephen D. Whitlow, Patricia May Ververs

Honeywell Laboratories
3660 Technology Drive; Minneapolis, MN 55418
michael.dorneich@honeywell.com

Abstract

This paper describes an evaluation of an adaptive system that significantly improved joint human-automation performance by "closing the loop" via utilization of a real-time, directly-sensed measure of cognitive state of the human operator. It is expected that in a highly networked environment the sheer magnitude of communication traffic amid multiple tasks could overwhelm the individual soldier. Key cognitive bottlenecks constrain information flow and the performance of decision-making, especially under stress. The Honeywell team, sponsored by the DARPA Augmented Cognition Program, is focusing primarily on varieties of the Attention Bottleneck: tonic arousal required for vigilance and divided attention across multiple tasks. Breakdowns in attention lead to multiple problems: failure to notice an event in the environment, failure to distribute attention across a space, failure to switch attention to highest priority information, or failure to monitor events over a sustained period of time. The appropriate allocation of attention is important to U.S. Army's Future Force Warrior (FFW) program because it directly affects two cornerstone technology thrusts within the FFW program: netted communications and collaborative situation awareness. Honeywell has developed a Closed Loop Integrated Prototype (CLIP) for application to FFW. The CLIP exploits real-time neurophysiological and physiological measurements of the human operator in order to create a cognitive state profile, which is used to trigger mitgation strategies to improve human-automation joint performance. The performance improvements of four mitigation strategies were studied in a Concept Validation Experiment (CVE): task management, modality management, task offloading, and task sharing. Findings indicate that performance can be improved by 100% or more by driving the mitigation strategies with knowledge about the operator's cognitive state.

1 Introduction

Honeywell is addressing the Attention Bottleneck in joint human-machine system performance. The Concept Validation Experiment (CVE) discussed in this paper is part of an on-going research (Dorneich et al 2003; Dorneich et al, 2004) effort to assess the viability of significantly improving joint human-automation performance by "closing the loop" via utilization of a real-time, directly-sensed measure of cognitive state of the human operator.

The U.S. Army is currently defining the roles of the 2010-era Future Force Warrior (FFW). The FFW program seeks to push information exchange requirements to the lowest levels and posits that with enhanced capabilities a squad can cover the battlefield in the same way that a platoon now does. Among other capabilities, the application of a full range of netted communications and collaborative situational awareness will afford the FFW unparalleled knowledge and expand the effect of the Future Force three dimensionally. Task analysis interviews with existing military operations identified factors that negatively impact communications efficacy. In the first few minutes of any intense mission, radio communications are a suboptimal method of communications because everybody is intensely focused on their tasks at hand. In one famous raid, for example, the commander did not hear the radio communications informing him that the plan had changed until he was physically grabbed by the ground force commander and given this critical information. The commander responded by radioing his own troops, who also did not respond. The implications of these kinds of situations are many, but, first and foremost, mission critical information must be reliably communicated.

Adaptive automation can either provide adaptive aiding, which makes a certain component of a task simpler, or can provide adaptive task allocation, which shifts an entire task from a larger multitask context to automation (Parasuraman, Mouloua, & Hilburn, 1999). Currently, adaptive systems derive their inferences about the cognitive state of the operator from mental models, performance on the task, or from external factors related directly to the task environment (Wickens & Hollands, 2000). For example, Scott (1999) used the projected time until impact as an external condition to infer that the pilot's attention was incapacitated, at which point the system would perform the "fly up" evasive maneuver to avoid a ground collision. In that case, the automation took over control of the system (i.e., the aircraft) from the pilot. Adaptive mitigation strategies include task management, optimizing information presentation via modality management, task sharing, and task loading. For instance, in task management, mitigation strategies might include intelligent interruption to improve limited working memory, attention management to improve focus during complex tasks, or cued memory retrieval to improve situational awareness and context recovery. For example, an air traffic controller might be presented with decision aids for conflict detection and resolution by the automated system when it detects a rapid increase in traffic density or complexity (Hilburn, Jorna, Byrne & Parasuraman, 1997). Modality management mitigation strategies might include utilizing available resources (i.e., audio, visual) to increase information throughput. Task offloading and task sharing to automation are also mitigation strategies to reduce workload. Ultimately, the goals of adaptive automation are similar to those of automation in general, such as avoiding "operator out of the loop" conflicts or mistrust in the automation.

An approach was adopted that considers the joint human-computer system when identifying bottlenecks to improve system performance (Dorneich et al, 2003). Key cognitive bottlenecks constrain information flow and the performance of decision-making, especially under stress. From an information-processing perspective, there is a limited amount of resources that can be applied to processing incoming information due to cognitive bottlenecks (Broadbent, 1958; Treisman, 1964; Kahneman, 1973; Pashler, 1994). The DARPA Augmented Cognition program identified four key cognitive challenges related to different components of information processing: 1) the sensory input bottleneck, 2) the attention bottleneck, 3) the working memory bottleneck, and 4) the executive function bottleneck (Raley, et al., in press). The Honeywell team, sponsored by the DARPA Augmented Cognition Program, is focusing primarily on the Attention Bottleneck, though the other bottlenecks are addressed in the studies described herein. There are many varieties of attention that need to be considered to optimize their distribution (Parasuraman & Davies, 1984): executive attention, divided attention, focused attention (both selective visual attention and selective auditory attention), and sustained attention. Breakdowns in attention lead to multiple problems: failure to notice events in the environment, failure to distribute attention across space, failure to switch attention to highest priority information, or failure to monitor events over a sustained period of time. The appropriate allocation of attention is important to FFW because it directly affects two cornerstone technology thrusts of the program: netted communications and collaborative situation awareness (Blackwell, 2003).

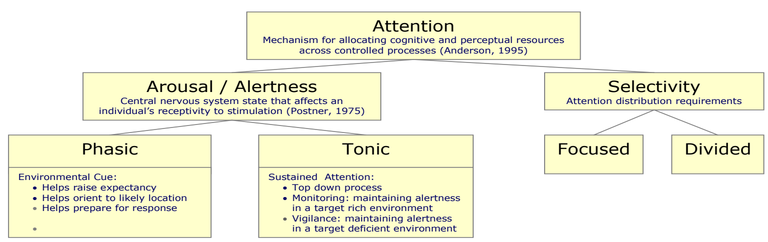

Figure 1. Simplified breakdown of Attention

Attention (see Figure 1) can be broadly defined as a mechanism for allocating cognitive and perceptual resources across controlled processes (Anderson, 1995). In order to perform effectively in military environments one must have the capacity to direct attention to task relevant events in a dynamic environment (*alertness*). Additionally, one must be able to narrow or broaden one's field of attention appropriately depending on the demands of a task (*selectivity*). Attention can be stimulated by external events (*phasic attention*) -- e.g. reacting to gunshots, or a loud aural warning. Alertness can also be maintained consciously, as a controlled top-down process (*tonic attention*). Examples of tonic attention include remaining vigilant while screening baggage at a security checkpoint, or looking for insurgents from surveillance positions over the span of hours. While phasic attention is mostly instinctive and automatic, tonic attention requires active effort on the part of a person. A vast body of literature attests to the difficulty of maintaining tonic attention over prolonged periods of time (e.g. Cabon et al, 1993, Colquhoun, 1985). The CVE explored the use of gauges to detect and drive mitigations during periods when tonic attention levels may be inadequate. We did so in the context of a vigilance task to be described later. Selectivity is another dimension of attention that is critical for task performance. Warfighters have to be able to distribute their attention over information sources effectively in order to accomplish various tasks. Attention has to be highly *focused* in many task contexts. Examples include a bomb disposal expert tuning out distractions to carry out intricate procedures associated with deactivating an incineary device, or a sniper taking aim at a target. However, many tasks require attention to be *divided* across a diverse range of information sources. This is particulary true in today's information centric warfare environment where the warfighter has to attend to potentially hostile events around him or her while maintaining communications and interacting with a range of information devices. An emphasis of the research reported here was a focus on performance under conditions where limited attentional resources have to be

4. Dorneich, M.C., Mathan, S., Creaser, J.I., Whitlow, S.D., & Ververs (2005). "Enabling Improved Performance though a Closed-Loop Adaptive System Driven by Real-Time Assessment of Cognitive State", *Foundations of Augmented Cognition* (D.D. Schmorrow, Ed.). Mahwah, NJ: Lawrence Erlbaum Associates. pp. 621-630.

distributed widely in order to perform effectively. Several of the experimental scenarios to be discussed later explored the efficacy of gauge driven mitigations under divided attention demands.

This paper describes an evaluation of the Augmented Cognition Adaptive system built by the Honeywell team. The adaptive system, described in section 2, is driven by real-time non-invasive neuro-physiological and physiological state detection techniques (known in this paper as "cognitive gauges") for determining cognitive workload and comprehension. Subject performed tasks in a virtual environment that represents dismounted soldier combat operations. This virtual environment allowed us to create, and test, a set of cognitive gauges. Cognitive workload was (broadly) defined as the amount of mental effort needed to perform satisfactorily on a task. Comprehension was defined as having sufficient cognitive resources to understand the information at the moment it is presented. Outputs of the cognitive gauges were used to drive an adaptive cognitive assistance system for dismounted combat operations. The experiment, described in section 3, looked at subject performance under a variety of attention states, and aided by a variety of mitigations strategies employed by the adaptive system. Results of the experiment is presented in section 4. With the aid of our proposed adaptive system we hope to increase the soldier's situation awareness, survivability, performance, and information intake by improving their ability to comprehend and act on available information. The results of this experiment will benefit soldiers by creating a system that alters task presentation based on an analysis of that soldier's cognitive state.

2 System Description

2.1 Architecture

The system architecture of the Honeywell Closed Loop Integrated Prototype (CLIP). (see Figure 2) is made up of the following components:

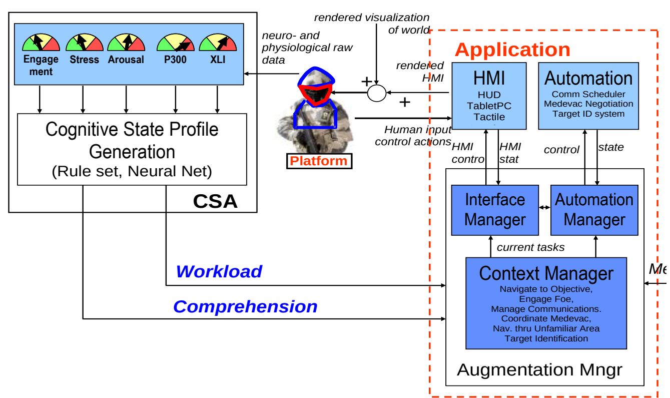

Figure 2. The CLIP architecture

- *Cognitive State Assessor* (CSA) combines measures of cognitive state to produces the cognitive state profile (CSP).
- *Human-Machine Interface*, where the human interacts with the system via a TabletPC, tactor belt, and radio.
- *Automation*, where tasks can be partially or wholly automated.
- *Augmentation Manager* (AM), adapts the work environment to optimize joint human-automation cognitive abilities for specific domain tasks, has:
  - *Interface Manager*, responsible for realizing a dynamic interaction design in the HMI
  - *Automation Manager*, responsible for the level and type of automation
  - *Context Manager*, responsible for tracking tasks, goals, and performance
- *Virtual Environment* (not shown) is a simulated approximation of the real world.
- *Experimenter's Interface* give the experimenter both insight and control over events within the system.

2.2 Cognitive State Assessor

2.2.1 Cognitive Guages

The Honeywell team has developed a comprehensive suite "cognitive gauges" for the CSA:.

- *Engagement Index*. The engagement index is a ratio of EEG power bands (beta/(alpha + theta)), and is a measurement of how cognitively engaged a person is in a task, or their level of alertness (Freeman et al., 1999). Adaptive systems have used this index to successfully control an automation system for tracking performance and a vigilance task (Freeman et al., 1999; Mikulka, Scerbo, & Freeman, 2002; Pope, Bogart, & Bartolome, 1995).The engagement index reflects the selection and focus on some aspect at the expense of the other competing demands, thus it is a measure of focused attention. Low levels of engagement indicate that the subject is not actively engaged with some aspect of the environment.

- *Arousal Meter*. Clemson University's Arousal Meter (Hoover & Muth, 2003) derives autonomic arousal from the cardiac inter-beat interval (IBI), derived from the Electrocardiogram (ECG) at one ms accuracy. The gauge has 3 levels (low, medium high). Increases in score are associated with increased autonomic arousal.
- *Stress Gauge*. The Institute of Human and Machine Cognition developed a composite Stress Gauge (Raj et al., 2003; Kass et al., 2003). The gauge uses a weighted average of the four inputs (Video Pupilometry (VOG), High Frequency Electrocardiogram (HFQRS ECG), Electrodermal Response (EDR), and Electromyogram (EMG) from the left trapezius muscle to detect the subject's response to changes in cognitive load. The gauge detects the cognitive stress related to managing multiple competing tasks on a moment-to-moment basis.
- *P300 Novelty Detector*. The P300-novelty detector gauge (Sjada, Gerson, & Parra, 2003), spatially integrates signals from sensors distributed across the scalp, to discriminate between task-relevant and task irrelevant responses. Mitigation strategies are based on the assumption that the presence of a strong evoked response indicates that subjects have sufficient attentional resources to process the incoming message..
- *Executive Load Index*. Human Bionics developed the eXecutive Load Index (XLI) (DuRousseau, 2004) operates by measuring power in the EEG at frontal (FCZ) and central midline (CPZ) sites. The index was designed to measure real-time changes in cognitive load related to the processing of messages.

2.2.2 Cognitive State Profile

The CSA outputs a Cognitive State Profile (CSP) comprised of two decision state variables: *workload* and *comprehension*. The CSP drives the mitigations of the Augmentation Manager. Currently a simple set of rules is used to derive the assessment of Workload and Comprehension, although work is underway to define this step with neural networks, in order to better account for individual subject differences (see companion paper in these proceedings: Mathan, Mazaeva, & Whitlow, 2005). For this CVE, *workload* was considered high is any of the three gauges, Engagement, Arousal, or Stress, registered high. This was done to allow for the fact that on any given subject, only a subset of the gauges may be able to discriminate differing levels of workload, based on individual differences in subject's cognitive response to the scenario workload manipulation. Likewise, in order to bias *comprehension* towards false positives, both the P300 Novelty Detector Gauge and the XLI gauge had to be high (i.e. that the subject *could* comprehend the message) for the *comprehension* variable to be set to "True".

2.3 Augmentation Manager

2.3.1 Communications Scheduler

Honeywell developed the Communications Scheduler to mitigate divided attention tasks via task-based management and modality-appropriate information presentation strategies.. The Communications Scheduler schedules and presents messages to the soldier based on the cognitive state profile (derived from the gauges), the message characteristics (principally priority), and the current context (tasks). Based on these inputs, the Communications Scheduler can pass through messages immediately, defer and schedule non-relevant or lower priority messages, escalate higher priority messages that were not attended to, divert attention to incoming higher-priority messages, change the modality of message presentation, or delete expired or obsolete messages. Messages are characterized by priority, depending on how critical they are: high (mission-critical and time-critical), medium (mission-critical only), and low (not critical). The Communications Scheduler determined the initial message presentation based on a user's current *workload*. After the first presentation of a message to the user (in audio modality), the Communications Scheduler determined whether to take further action on a message depending on the CSA's assessment of *comprehension*. A more detailed description of the Communications Scheduler can be found in the companion paper in these proceedings (Dorneich et al 2005).

2.3.2 Tactile Navigation Cueing System

The Tactile Situation Awareness System (TSAS) provides navigation cueing during mitigated trials, and consists of a 12 pairs of tractors, representing the cardinal positions of the clock (12 o'clock centered on the umbilicus), worn about the upper abdomen. Operationally, pulses from tectors "tug" the subjects in direction to go. The system was invoked when the CSP indicated *workload* was high and the subject needed to navigate through an unfamiliar route.

However, turning the system off as soon as *workload* fell below some threshold would leave users disoriented in an unfamiliar area. Thus the navigation mitigation persists until users get to the safe destination.

2.3.3 Medevac Negotiation Agent

The evacuation of injured personnel is a crucial warfighter function. The task is lengthy and requires a substantial amount of information to be communicated accurately. Performance on this task may suffer under high workload conditions. Personnel may omit important information or make errors in the information transmitted. Additionally, attention devoted to coordinating the medevac exchange may detract from the performance of other critical tasks. The Medevac Negotiation Agent was triggered on if *workload* was high and a medevac had to be coordinated. A medevac icon on the HUD notified the user about the need to coordinate an evacuation using the medevac agent. Information available on the FFW Netted Communications network would be automatically filled in, and the system would present this information about casualties to the platoon leader for inspection. The platoon leader could review the medevac information on the Tablet PC and transmit information using the interactive form. The system also provided the option to delegate subsequent medevac negotiation to team members facing lower workload demands.. The medevac agent only contained the most critical information needed for a Medevac. A more detailed information exchange might allow for safer and more efficient Medevac operations. Additionally, engaging in Medevac communications may contribute to better situational awareness of a team's status. It is for these reasons that the augmentation is only invoked when the subject's workload is so high and the subject's performance so inadequate, that the differences associated with automated negotiation are acceptable in terms of overall performance.

2.3.4 Mixed-Initiative Identification System

Military personnel sometimes have to maintain high levels of sustained attention in environments where target stimuli may be infrequent and hard to detect. Research suggests that performance on these tasks deteriorates considerably over time. This task cannot be easily automated, since automated systems trained to detect target stimuli in a field may not perform as well as an alert human. However, these systems could play a helpful role if they can be triggered when gauges detect a vigilance decrement. Subjects in the CVE lookined for targets in a series of surveillance photos. If Engagement was low, the system assists in identifying targets by drawing a yellow box around any detected targets. This system was modeled after enabling technology (Schneiderman & Kanade, 2004) currently under development. Such a system of mixed initiative search with intelligent assistance is part of the FFW vision (US Army, 2003).Due to a non-acceptable frequency of errors within the system, the subject used system output for advice, but must continue to scan. In addition, other issues such as over-reliance on automation and the human's generally poor ability to passively supervise automated processes preclude automating the process entirely.

2.4 Virtual Environment

The operational environment was a desktop first-person virtual environment (VE) that simulates Mobile Operations in an Urban Environment . The VE consisted of a city of narrow streets, surrounded by two and three story buildings. The environment had an industrial appearance. The visual complexity of the environment contributed to the subjects' workload. The subject was faced with some number of enemy forces, presented both at street level and above. The enemy forces had attacked with varying levels of success (depending on the workload and difficulty settings). The subject performed all tasks in the environment using a combination of keyboard and mouse controls. The controls allowed the subject to look around the world, to move (walking forward or backward, sidestepping left or right, jumping, and crouching), to shoot their weapon (an approximation of an M16) and to manage messages.

3 Experimental Method

Can the gauges detect the cognitive states of interest consistently enough to drive the mitigations? If the mitigations are driven correctly, do we see performance improvements on the communications? The goal for the CVE was to improve performance on mitigated tasks, with no performance decrement to concurrent tasks, and with no negative effect on overall workload. Three scenarios were developed for the CVE, where each scenario looked at different states of attention, tasks, and mitigations. The CVE focused on two types of attention: Tonic arousal required for vigilance (Scenario 3) and divided attention across multiple tasks (Scenarios 1 and 2).

4. Dorneich, M.C., Mathan, S., Creaser, J.I., Whitlow, S.D., & Ververs (2005). "Enabling Improved Performance though a Closed-Loop Adaptive System Driven by Real-Time Assessment of Cognitive State", *Foundations of Augmented Cognition* (D.D. Schmorrow, Ed.). Mahwah, NJ: Lawrence Erlbaum Associates. pp. 621-630.

3.1 Operational Scenarios

3.1.1 Scenario 1: Divided Attention

Scenario 1 focused on three critical task elements: "Navigate to Objective," "Engage Foes" and "Manage Communications" The subject ,a platoon leader, led his or her platoon along a known route through a hostile urban environment to the objective, while being careful to engage enemy soldiers. Subjects also received incoming communications throughout, some requiring an overt response. Subjects received status and mission updates, requests for information, and reports; these communications are a primary source of their situation awareness. Radio communications volume was extremely high. The scenario only included two or three high priority messages (hold at certain locations for a specified amount of time, or objective location had changed). Failure to heed these high priority messages caused the subject to encounter an ambush. Figure 3 details the route, the points in the scenario where the high priority messages occurred, and the potential ambush locations, if should the subject fails to heed the messages. Scenario 1 was designed to put subjects into distinct periods or extreme high and extreme low workload periods. Validation of the ability of the gauges to distinguish between these periods is vital. When they held at a location for up to three minutes, a low task load situation, it was expected that the gauges would register low workload. When subjects engaged in a firefight, the gauges were expected to register cognitive states associated with high workload. Scenario 1 was principally designed to test the performance improvements derived from the Communications Scheduler. Thus the high workload times included a high volume of communications traffic to the subject, just at the time when their workload was high due to being targeted by foes. Metrics are defined for each task in Table 1.

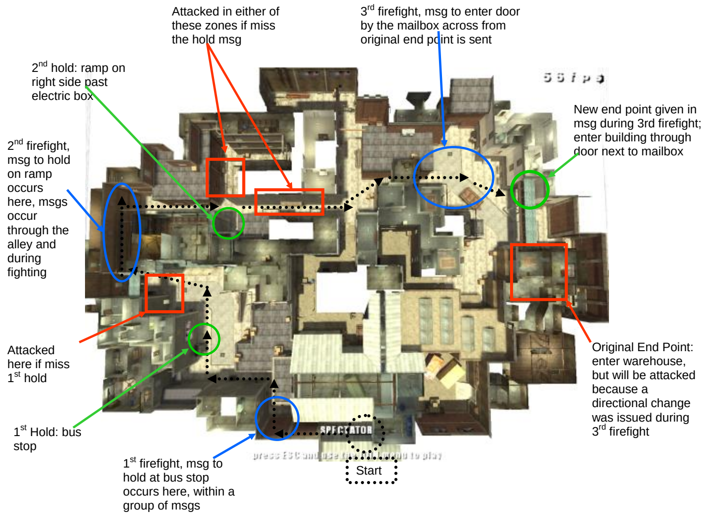

Figure 3. Scenario 1: Divided Attention

Table 1. Task metric definitions for Scenario 1

Metric	Definition	Task
Message comprehension	Message comprehension is inferred from compliance with the orders or correct responses from direct quieires in messages.	Manage Communications
Situation Awareness	Score on 4-8 questions asked after each scenario to ascertain if they could recall mission-critical information relayed through the communications.	Manage Communications
Run Time	Time subjects took to complete a known (trained) route, while avoiding hits, was a measure of their effectiveness to navigate to the objective.	Navigate to Objective
Hits Taken	Number of hits by OPFOR (Opposing Forces) on subject.	Engage foes
Hits on OPFOR	Number of hits by subject on OPFOR.	Engage foes
Shot Accuracy	The percentage of shots fired that hit OPFOR	Engage foes

3.1.2 Scenario 2: Divided Attention

The subject traversed the same initial route as in Scenario 1. However, upon reaching the objective area, the subject was informed that the enemy had set a trap. The subject needed to abandon the objective, get back to the safe zone, and avoid the route the just taken to the objective. To return to the safe zone, the subject had to navigate though unfamiliar parts of the city in order to avoid ambushes. The task load stems from having to mentally convert an exocentric 2D representation of an unfamiliar area into an ego centric representation and reasoning with this newly formed representation. The participants were provided with an updated map on their Tablet PC showing

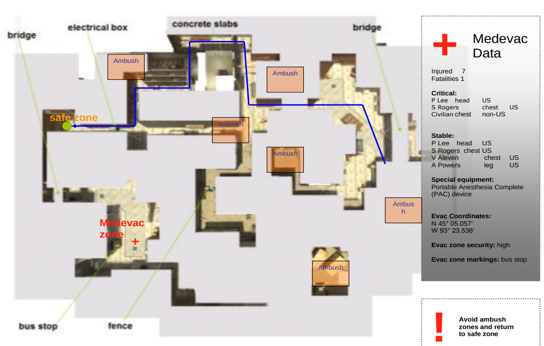

Figure 4. Scenario 2: Map and MEDEVAC

potential ambush zones. Simultaneously, the participants got a request to coordinate a MEDEVAC immediately. This is a quite lengthy and communications- and information-intensive procedure. The map and the information requirements of the MEDEVAC procedure are illustrated in Figure 4. Both the Medevac communications procedure and the navigation to the safe zone task had to be accomplished simultaneously and under extreme time pressure. Additionally the subject had to engage any foes encountered. Metrics are defined for each task in Table 2.

Table 2. Task metric definitions for Scenario 2

Metric	Definition	Task
Time to Safe Zone	Time to navigate from the warehouse to the safe zone	Nav. thru Unfamiliar Area
Ambushes Encountered	Number of ambushes encountered..	Nav. thru Unfamiliar Area
Medevac Questions Answered	Number of medevac-related questions answered correctly	Coordinate Medevac
Time to complete Medevac	Time to complete the medevac negotiation process	Coordinate Medevac
Hits Taken	Number of hits by OPFOR on subject.	Engage foes
Hits on OPFOR	Number of hits by subject on OPFOR.	Engage foes

3.1.3 Scenario 3: Sustained Attention

In Scenario 3, the subject was the leader of a recon unit, and was responsible for identifying any targets (enemy soldiers) from reconnaissance photos received (updates every 2 seconds), via his or her Tablet PC, from external surveillance cameras. The experimental protocol for this scenario was a classic vigilance paradigm. The scenario lasted approximately 30 minutes. The first five minute session had target occurring at a rate of 14%, and served as the measure of baseline performance. This was followed by a 20-minute session with a very low target occurrence rate (3%), to produce a vigilance decrement. The final five minute session had a target occurrence rate identical to the first five minute session. For Scenario 3, the metric of interest is:

- *Target Identification*. The accuracy to which subjects identified targets was taken as a measure of performance. Specifically, the accuracy of stage 1 (baseline) was compared with performance in Stage 3.

3.2 Mitigation Strategies Applied

There are four broad categories of possible mitigations in an Augmented Cognition system: 1) task/Information management, 2) modality management, 3) task offloading, and 4) task sharing. The multiple scenarios of the IHMC CVE provided Honeywell with the opportunity to explore a wide range of possible mitigation strategies. All four categories of mitigation strategies are addressed by the Augmentation Manager. Table 3 summarizes the Mitigation Agents that realize classes of mitigation strategies..

Table 3. Summary of mitigation agents that realize classes of mitigation strategies.

Mitigation Strategy	Mitigation Agent	Task Mitigated	Scenario
Task Scheduling	Communications Scheduler	Manage Communications	1
Task Offloading	Medevac Negotiation	Coordinate Medevac	2
Task Sharing	Mixed-Initiative Target ID	Target Identification	3
Modality Mgmt	Communications Scheduler	Manage Communications	1
Modality Mgmt	Tactile Navigation Cueing	Navigate thru Unfamiliar Area	2

While the mitigations described here have the potential for boosting performance when human cognitive resources may be limited, they could have detrimental effects if left on at all times. Gauge driven mitigation allows these mitigations to be activated when the benefits are likely to outweigh the costs. See the companion paper (Mathan, Dorneich, & Whitlow, 2005) in the proceedings for a thorough discussion of the role of automation etiquette in the design of effective cognitive-state-driven mitigation strategies.

3.3 Experimental Design

The CVE consisted of three two-factor experiments. Each experiment compared performance under gauge driven mitigation with performance without mitigation. For scenarios 1 and 2 this evaluation was a within subjects design, as each subject saw both scenarios in both the mitigated and unmitigated cases. Half of the subject saw Scenario 3 with mitigation, and the half without. Thus for scenario 3 the experiment is a between subjects design.

Hypothesis. In general, the hypothesis for this experiment is as follows: the mitigations will improve performance on the tasks they are mitigating without decrementing other concurrent tasks**.**

Participants. 14 males (M_{age} = 25.4 years, avg. education = 15 years) participated in this experiment. To reduce the effect of learning for this experiment, participants were chosen who rated their skill level (avg = 3.4 of 5-scale, avg playing time = 5.7 hours/week) at first person shooter games as average to above average. .

Independent Variables. 1) Mitigation (on/off), and 2) Scenario (three, which vary by attention type).

Dependant Variables in addition to metrics defined for each scenario (e.g. see Table 1),subjective workload (via NASA TLX) and subject preferences were assessed.

4 Results

4.1 Scenario 1 Task Metrics

This scenario focused on divided attention bottleneck in multi-tasking, and consisted of the subject performing three tasks. The mitigation strategy employed in this scenario was the Communications Scheduler. The goal was to improve performance on the Manage COmmunciations task while not decrementing performance on the Navigate to Objective and Engage Foes task.

4.1.1 Manage Communications Task Metrics

Message comprehension. Subjects in the unmitigated condition correctly responded to 57 of 143 possible messages (39.9%); . in the mitigated condition they correctly responded to 114 of 143 messages (79.7%). The mitigated condition shows a significant (p<0.0001) performance increase of 100% (see Figure 5).

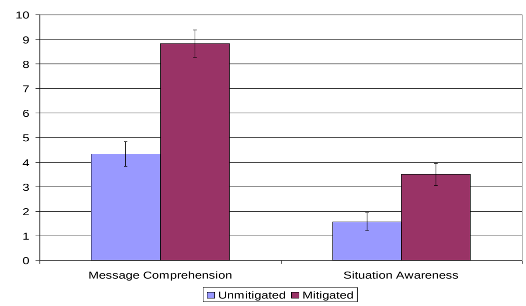

Figure 5. Manage Communications task

Situation Awareness. Subjects in the unmitigated condition correctly responded to 22 of 84 situation awareness questions (26.2%). Subjects in the mitigated condition correctly responded to 49 of 84 situation awareness questions (58.9%). The mitigated condition shows a significant (p=0.009) performance increase of 125% (see Figure 5).

4.1.2 Navigate to Objective Task Metrics

Run Time. The total run time for the mitigated condition (M = 965 sec) was significantly longer than the run time for the unmitigated condition (U = 469 sec), t(12) = -8.29, p < 0.001. This result is in the expected direction. In the mitigated condition, more participants heard the hold messages, which increased the time mission completion time.

4.1.3 Engage Foes Task Metrics

Hits Taken. There was no significant performance change for hits taken while engaging foes (p = 0.29).

Hits on OPFOR. There was a significant difference between the unmitigated (M = 518 sec) versus the mitigated (M = 468) for how many times the subject was able to shoot opposing forces, $t(12) = 2.93$, $p = 0.013$. Participants shot the opposing force significantly fewer times during the mitigated condition. This result is expected because in the mitigated condition subjects comprehended the “hold” message more often and thus avoided ambushes.

Shooting Accuracy. No significant performance change for shooting accuracy while engaging foes (p = 0.06). .

4. Dorneich, M.C., Mathan, S., Creaser, J.I., Whitlow, S.D., & Ververs (2005). "Enabling Improved Performance though a Closed-Loop Adaptive System Driven by Real-Time Assessment of Cognitive State*", Foundations of Augmented Cognition* (D.D. Schmorrow, Ed.). Mahwah, NJ: Lawrence Erlbaum Associates. pp. 621-630.

4.2 Scenario 2 Task Metrics

This scenario focused on divided attention bottleneck in multi-tasking. The mitigation strategies employed in this scenario were the Tactile Navigation Cueing System and the Medevac Negotiation Tool. These mitigations were expected to improve performance on the Navigate Through Unfamiliar Area task and Coordinate Medevacn task, respectively, all while not degrading performance on the Engage Foes task.

4.2.1 Navigate Through Unfamiliar Area Task Metrics

Time to Safe Zone. The total time it took a subject to reach the safe zone was greater for unmitigated subjects than mitigated subjects. The data is illustrated in [create new figure?]. The average performance improvement was 20%, however this difference was not significant (p = 0.30). [need numbers]

Ambushes Encountered. Subjects in the unmitigated case were almost four times as likely to navigate into an ambush than subjects in the mitigated case. Unmitigated subjects (N=12) ran into 19 ambushes, while the mitigated subjects (N=12) ran into 5 ambushes. The difference was significant ($p<0.003$). The mitigation resulted in a 380% performance improvement.

4.2.2 Coordinate Medevac Task Metrics

Medevac Questions Answered. Subjects in the unmitigated case answered 50 questions correctly out of a possible 98 (51% correct). Subjects in the mitigated case answered 98 of 98 questions correctly (100% correct). Thus the mitigation was able to significantly ($p=0.004$) increase performance by 95%, as shown in Figure 6.

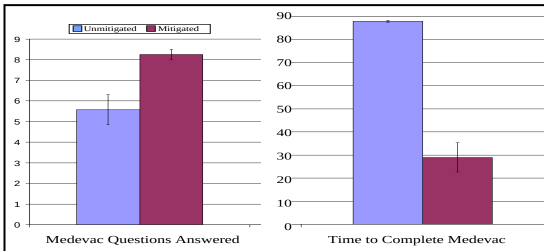

Figure 6. Coordinate Medevac task metrics.

Time to Complete Medevac. Subjects where able to complete the medevac task significantly faster in the mitigated condition ($p < 0.001$), resulting in a 303% performance improvement, as shown in Figure 6.

4.2.3 Engage Foes Task Metrics

Hits Taken and Hits on OPFOR.,Participants took fewer hits from the OPFOR during the mitigated (M = 102) than the unmitigated (130) condition, $t(12) = 2.42$, $p = 0.03$. Additionally, participants had fewer hits on OPFOR in the mitigated (M = 22) versus the unmitigated (M = 27) condition, $t(12) = 2.47$, $p = 0.029$. Participants in the mitigated condition potentially encountered fewer ambushes because their ability to navigate the safe route was improved , resulting in participants seeing fewer enemy forces and thus, receiving and inflicting fewer hits.

4.3 Scenario 3 Task Metrics

Target Identification. Baseline performance (stage 1, no mitigation) was 65.8%. Stage Stage 3 mitigation was set at an accuracy rating of 68%. Unmitigated subjects in Stage 3 had an accuracy of 66.2%. Thus, on average, the experiment was not able to produce the decremented human performance one like to see in a vigilance experiment. Nonetheless, subjects in the mitigated condition performed at an accuracy of 85%, much better than the human (66.2%) or automation (68%) accuracy alone. The 30% performance improvement was significant ($p=0.022$).

4.4 Subjective Workload and Preferences

None of the scales nor the total overall workload were significantly different for either Scenario 1 or 2. Overall 76.9% of the subjects thought it was easier to perform tasks in the mitigated condition. Subjects also found mitigated condition easier for fighting (61.5%), communicating (84.6%) and navigating (76.9%).

5 Acknowledgements

This paper was supported by a contract with DARPA and the U.S. Army Natick Soldier Center, DAAD-16-03-C-0054, for which LCDR Dylan Schmorrow serves as the Program Manager of the DARPA Augmented Cognition program and Mr. Henry Girolamo is the DARPA Agent. The opinions expressed herein are those of the authors and do not necessarily reflect the views of DARPA or the U.S. Army Natick Soldier Center.

The authors would like to acknowledge the contributions of Mr. Jim Carciofini, Mr. Trent Reusser, and Mr. Jeff Rye, and Ms. Danni Bayn of Honeywell. Additional thanks to the Institute for Human and Machine Cognition, lead by Dr. Anil Raj, for hosting the experiment: Mr. Roger Carff, Mr. Matt Johnson, and Mr. Jeremy Higgins .

References

Anderson, JR (1995). Cognitive Psychology and its Implications (2nd Ed.). New York: Freeman

Blackwell, C. (2003), "Objective Force Warrior: Advanced Technology Demonstration", *Natick Soldier Center*, presentation to Honeywell at the NATICK IPR, October 23.

Broadbent, D. E. (1958). *Perception and Communication*. New York: Pergamon.

Cabon, Ph., Coblentz, A., Mollard, R. & Fouillot, J.P. (1993) 'Human Vigilance in Railway and Long-Haul Flight Operation', Ergonomics, 36(9): 1019-1033.

Colquhoun, W.P. (1985) 'Hours of Work at Sea: Watchkeeping Schedules, Circadian Rhythms and Efficiency', Ergonomics, 28(4): 637-653.

Dorneich, M.C., Whitlow, S.D., Mathan, S., Carciofini, J., and Ververs, P.M. (2005). "The Communications Scheduler: A Task Scheduling Mitigation For A Closed Loop Adaptive System." *Proceedings of the11th International Conference on Human-Computer Interaction (Augmented Cognition International)*, Las Vegas, Jul 22-27, 2005

Dorneich, M., Whitlow, S., Ververs, P.M., Carciofini, J., and Creaser, J. (2004) "Closing the Loop of an Adaptive System with Cognitive State," *Proceedings of the Human Factors and Ergonomics Society Conference 2004*, New Orleans, September.

Dorneich, Michael C., Whitlow, Stephen D., Ververs, Patricia May, Rogers, William H. (2003). "Mitigating Cognitive Bottlenecks via an Augmented Cognition Adaptive System," *Proceedings of the 2003 IEEE International Conference on Systems, Man, and Cybernetics*, Washington DC, October 5-8, 2003.

DuRousseau, D.R., (2004). *Multimodal Cognitive Assessment System*. Final Technical Report, DARPA, DAAH01-03-C-R232.

Freeman, F.G., Mikulka, P.J., Prinzel, L.J., & Scerbo, M.W. (1999). Evaluation of an adaptive automation system using three EEG indices with a visual tracking system. *Biological Psychology*, *50*, 61-76.

Hoover, A. & Muth, E.(in press). A Real-Time Index of Vagal Activity. *International Journal of Human-Computer Interaction*.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kass, S J, Doyle, M Raj, AK, Andrasik, F, & Higgins, J (2003, April). Intelligent adaptive automation for safer work environments. In J.C. Wallace & G. Chen, *Occupational health and safety: Encompassing personality, emotion, teams, and automation*. Symposium conducted at the Society for Industrial and Organizational Psychology 18th Annual Conference, Orlando, FL.

Mathan, S., Mazaeva, N., & Whitlow, S. (2005). Sensor-based cognitive state assessment in a mobile environment. *Proceedings of the 11th International Conference on Human-Computer Interaction (Augmented Cognition International)*, Las Vegas, Jul 22-27, 2005.

Mathan, S., Dorneich, M., & Whitlow, S. (2005). Automation Etiquette in the Augmented Cognition Context. *Proceedings of the 11th International Conference on Human-Computer Interaction (Augmented Cognition International)*, Las Vegas, Jul 22-27, 2005.

Mikulka, P. J., Scerbo, M. W., & Freeman, F. G. (2002). Effects of a Biocybernetic System on Vigilance Performance. *Human Factors*, *44*(4), 654-664.

Parasuraman, R. & Davies, D. R. (1984). *Varieties of Attention*. New York: Academic Press.

Pashler, H. (1994) Dual-task interference in simple tasks: data and theory. *Psychological Bulletin, 116*, 220-244.

Pope, A.T., Bogart, E.H., & Bartolome, D.S (1995). Biocybernetic system validates index of operator engagement in automated task. *Biological Psychology, 40*, 187-195.

Prinzel, III, L.J., Hadley, G., Freeman, F.G., & Mikulka, P.J. (1999). Behavioral, subjective, and psychophysiological correlates of various schedules of short-cycle automation. In M. Scerbo & K. Krahl (Eds.), *Automation technology and human performance: Current research and trends*. Mahwah, NJ: Lawrence Erlbaum Associates.

Raj A.K., Perry, J.F., Abraham, L.J., Rupert A.H. (2003) Tactile interfaces for decision making support under high workload conditions, *Aerospace Medical Association 74th Annual Scientific Meeting*, San Antonio, TX..

Raj A.K., Kass, S.J., and Perry, J.F., (2000). "Vibrotactile displays for improving spatial awareness". In the *Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting*. Santa Monica, CA: HFES.

Raley, C., Stripling, R., Schmorrow, D., Patrey, J., & Kruse, A. (2004). Augmented Cognition Overview: Improving Information Intake Under Stress. In the *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*. Santa Monica, CA: HFES.

Schneiderman, H. and Kanade, T. (2004) "Object Recognition Using Statistical Modeling". http://www.ri.cmu.edu/projects/project_320.html.

Sajda, P., Gerson, A., & Parra, L. (2003). Spatial Signatures of Visual Object Recognition Events Learned from Single-trial Analysis of EEG, *IEEE Engineering in Medicine and Biology Annual Meeting*, Cancun, Mexico.

Treisman, A.M. (1964). Verbal cues, language, and meaning in selective attention. *American Journal of Psychology, 77*, 206-219.

US Army (2003), "*The Objective Force Warrior: The Art of the Possible…a Vision*," Available Online at http://www.ornl.gov/sci/nsd/pdf/OFW_composite_vision.pdf.