12TH ICCRTS
"Adapting C2 to the 21st Century"

Title

Using Physiological Sensors to Understand, Measure and Adapt to Stressors in the C2 Environment

Topics

C2 Metrics and Assessment, Cognitive and Social Issues, C2 Technologies and Systems

Authors

Jennifer McKneely (POC)
Johns Hopkins University Applied Physics Lab
MS 7-354
11100 Johns Hopkins Rd.
Laurel, MD 20723
240-228-7476
jennifer.mckneely@jhuapl.edu

Kevin Cropper
Johns Hopkins University Applied Physics Lab
MS 7-354
11100 Johns Hopkins Rd.
Laurel, MD 20723
240-228-7967
kevin.cropper@jhuapl.edu

William Fitzpatrick
Johns Hopkins University Applied Physics Lab
MS 7-354
11100 Johns Hopkins Rd.
Laurel, MD 20723
240-228-1382
william.fitzpatrick@jhuapl.edu

Jennifer Ockerman, PhD
Johns Hopkins University Applied Physics Lab
MS 7-354
11100 Johns Hopkins Rd.
Laurel, MD 20723
240-228-0764
jennifer.ockerman@jhuapl.edu

Nathan Koterba
Johns Hopkins University Applied Physics Lab
MS 7-354
11100 Johns Hopkins Rd.
Laurel, MD 20723
240-228-5947
nathan.koterba@jhuapl.edu

Leigh Baumgart
Johns Hopkins University Applied Physics Lab
MS 7-354
11100 Johns Hopkins Rd.
Laurel, MD 20723
240-228-2440
leigh.baumgart@jhuapl.edu

Abstract

Stress and fatigue will always be present in command and control centers. Understanding how stress and fatigue affect decision making performance will provide essential information needed to design and evaluate advanced command and control (C2) systems and environments. The ability to objectively measure cognitive attributes, such as workload or situation awareness, would provide detailed knowledge of how physiological measurements relate to general performance, decision-making, and situational awareness and would greatly enhance our ability to evaluate current and proposed command and control systems as well as design new systems that optimize cognitive attributes and capabilities during both operations and training.

Johns Hopkins University Applied Physics Laboratory (JHU/APL) is setting up a physiologic sensor suite to investigate the relationships between the physiological measures provided by the sensors and a warfighter's performance under various conditions of stress and fatigue. The sensors will include an electroencephalogram (EEG), an electrocardiogram (ECG), blood pressure monitor, temperature, and respiratory monitors, a galvanic skin response (GSR) sensor, mobile and stationary eye trackers, and software to combine and integrate the sensor outputs. This sensor suite will be integrated into a simulated C2 environment at JHU/APL.

One goal will be the development of *objective, non-intrusive* human performance measurement techniques using a physiological sensor suite to uncover the relationships between warfighter cognitive state, physiological measures, and warfighter performance. This work will build on previous research by the Augmented Cognition community and internal research completed by the first author investigating the relationship between fatigue and situation awareness.

This paper provides a background into the discipline of Augmented Cognition, a description of the physiological suite that JHU/APL is installing, a description of past research into fatigue and its relation to situation awareness, and a discussion on future activities to be pursued.

Introduction

Stress and fatigue are pervasive in modern life, and military life is no exception. Stress and fatigue will always be issues in command and control centers and can significantly affect military operations. Stress and fatigue alter warfighters' cognitive states, which affects their ability to perceive, comprehend, and understand incoming data and information, as well as impact their decision making abilities.

Without an ability to objectively and unobtrusively collect data on a warfighter's cognitive state, it is impossible to evaluate new and existing command and control systems for operational effectiveness and impact on a warfighter's cognitive state and decision making performance. No current methods exist to objectively and non-intrusively measure cognitive attributes, such as workload or situation awareness. Such ability would greatly enhance our capability to evaluate current and proposed command and control systems as well as design new systems that optimize a warfighter's current capabilities.

One domain investigating these issues is augmented cognition. Augmented cognition started as a DARPA initiative in 2001, though the term found broad use in 2000.

Augmented Cognition Background

Augmented cognition strives to mitigate the limitations and extend (augment) the capabilities of the human mind to greatly increase an operator's ability to perceive, comprehend and understand incoming data and information. The principal concept enabler is the development of technology that adapts the operator's environment to the operator's current mental state, which can change quickly over time due to factors such as stress (from boredom to over-excitement) and fatigue. Thus, two broad research areas exist in augmented cognition: 1) developing methods to objectively measure an operator's cognitive state and 2) developing methods of altering an operator's environment to account for their current cognitive state.

In the first research area -- developing methods to objectively measure an operator's cognitive state -- many techniques are being tried. These include the use of Electroencephalograms (EEG) readings of various brain waves and electrical impulses, Electrocardiogram (ECG) readings, respiration rates and depths, galvanic skin responses, and eye tracking. In addition, a significant amount of work exists looking at filtering each of these sources of data for useful signals and fusing data from multiple sources together. This work requires an interdisciplinary team from fields such as neuroscience, physiology, biopsychology, cognitive science, computer science, human factors and information management.

In the second research area – developing methods of altering an operator's environment to account for their current cognitive state – the focus has been primarily on altering the information content and display. For example, when a warfighter is perceived to be stressed the system may only provide the information required for immediate use, as opposed to a more comprehensive set of information.

[More detailed background information with references will be provided in the final paper.]

JHU/APL Physiological Suite

The JHU/APL Physiological Suite consists of several different products from various vendors as shown in Table 1. The decision of which equipment to purchase was based on an evaluation of cost, ease-of-use, reputation, and previous use in the Augmented Cognition research community. The equipment is currently located in a JHU/APL command and control laboratory, the Precision Engagement Transformation Center (PETC). However, since the equipment is light, portable, and packaged in carrying cases, it can easily be relocated to other research locations.

Two laptops provide the necessary computing for the suite. One laptop is dedicated to the Seeing Machines stationary eye tracker. The other laptop is used as a processing unit for the mobile eye tracker and collects data from the biofeedback sensor suite developed by Thought Technology.

Table 1: JHU/APL Physiological Suite

Device		**Measures (Physiological component)**	**Equipment**	**Image**
Eye Tracker	Head-Mounted	Gaze, PERCLOS	Applied Science Laboratories Mobile Eye (Lightweight Tetherless Eye Tracking)	
	Stationary	Gaze, PERCLOSE	Seeing Machines FaceLab 4.0 with Analog Stereo Tracking Cameras	
Skin Response Sensor		Galvanic Skin Response	Thought Technology GSR/TEMP 2X	
Biofeedback Sensor Suite		ECG, Temperature, Respiration, Heart rate/Blood pressure	Thought Technology ProComp Infiniti	
EEG Sensor Suite		EEG	Still to be purchased	

Fatigue and Situation Awareness

Human decision making and situation awareness (SA) are increasingly critical components to warfighting effectiveness. Situation awareness or the ability to understand what is happening in your surroundings, how things are changing, and predicting what could happen in the future can be hampered by large amounts of data and information generated by automated systems that an individual must parse through. To ensure maximum effectiveness and ability to maintain SA, warfighters must remain alert and cognizant of the operating environment around them.

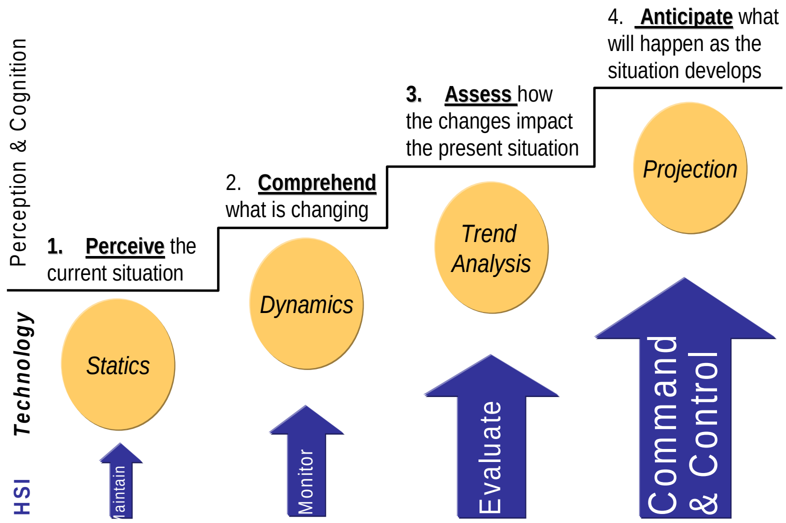

Figure 1. Four Level SA Definition

In that situation awareness is foundational to decision making, it is a useful human performance metric to include in the evaluation of systems. Decision-making and situation awareness become even more critical in today's varying operational environments which are employing fewer people who are expected to perform at higher levels of efficiency than ever before. JHU/APL developed a SA definition and measurement tool (Provisional Patent 2198-6606). The SA definition is based on Endsley's (2000) definition of SA in which level 1 is perception, level 2 is comprehension and level 3 is projection; however, we have added a level between Endsley's level 2 and level 3 for trend analysis (see Figure 1). Trend analysis enables the tracking of how things change over time assisting in the prediction of what will happen next. Hence, the SA definition is separated into levels in which level 1 was perception, level 2 was comprehension, level 3 was trend analysis, and level 4 was projection.

Even though numerous SA measures exist and have been used in military applications, they often include significant limitations. For example, some require briefly stopping the warfighter's activity to answer questions (e.g. Situation Awareness Global Assessment Technique, SAGAT) or are collected at the end of a task (e.g. Situational Awareness Rating Technique, SART) and provide only subjective measures. In addition, these SA metrics often do not result in measures that can be easily related and compared to other metrics. Therefore, an SA assessment method has been developed and is complemented with a novel metric tool that can be used in real-time. This comprehensive and flexible tool increases the integrity of SA data collection, resulting in a single weighted score. A useful function of our proposed SA metric is that it can be correlated with other metrics to evaluate relationships and interactions between various factors.

In a series of studies, JHU/APL has demonstrated the ability to measure situation awareness over extended durations and tested the new SA principle, process, and measurement. In addition the relationship between fatigue (as measured with physiological sensors and reaction time tests) and situation awareness was investigated. A framework to measure SA in real-time was developed through research into the state-of-the-art in situation awareness, measurement of SA, and its relationship with fatigue.

To measure SA, SA assessment/probe questions applicable to the operational task were developed and assigned a level of SA (selected examples from the second study investigating Undersea Warfare are listed in Table 2). Delivery and recording of probes was executed in a realistic manner with a SME role playing supervisors seeking information that would normally be requested. Responses were recorded to be scored after the conclusion of the experiment. While participants performed the SA task, the test administrator recorded the participants' responses and judged the participants' confidence. Participants also subjectively assessed their confidence with the modified version of the Situational Awareness Rating Technique (SART). The SART typically includes ratings of the supply and demand of attention and understanding; the modified version also included a rating of self-confidence.

Table 2. Sample SA questions at the four levels of SA.

Question	SA Level
1. Are there any active contacts?	1
2. What was the last bearing?	1
3. Is the towed array stable?	2
4. Which contacts would you classify as non-sub?	2
5. Are contacts getting weaker?	3
6. What way would you turn to resolve bearing ambiguity?	3
7. What maneuver would you recommend to regain contact?	4
8. Do you predict that target strength will decrease?	4

It is hypothesized that situation awareness decreases as fatigue increases, and that using more of our senses may offset the effects of fatigue. As a person becomes fatigued, different cognitive capabilities degrade at different rates. Current fatigue detectors are validated for detecting Stage 0, the onset of sleep. However, experience shows that cognitive function starts degrading before this point, and additionally, potential countermeasures may take a period of time to restore cognitive effectiveness. The operational impact of the time period between fatigue-induced cognitive degradation, fatigue detection, and countermeasure effectiveness is uncertain.

At this point in time, the causal links between C2 tasks, their underlying cognitive functions, fatigue sensor capabilities and countermeasure implementation are not well understood. Once these relationships are understood and fatigue can be detected early, countermeasures can be developed to prevent performance degradation. Further when the relationship between situation awareness and fatigue is fully understood, it is possible to detect fatigue by assessing situation awareness. A non-invasive, objective method of measuring SA is still needed. The field of cognitive neuroscience is progressing to a point where new technologies and scientific insights are enabling measurement of cognitive state. This may provide the capability to uncover physiological patterns related to SA, similar to acquisition of skill. It has been found that there are brain patterns associated with errors such that differentiation between a slip (incurred unintentional action) and a mistake (incorrect intentional action) can be made (Luu and Campbell, 2005). Results from these studies provide a foundation for future efforts investigating designs to support SA in fatigued states for command and control.

Results of the first, pilot study (McKneely, Bevan, Cropper, Iny, & Vaughan, 2005) supported the use of the SA measurement tool to assess fatigue effects on SA and the use of physiological measures to measure cognitive task performance. Results were consistent with previous literature supporting the correlation between established fatigue tests and biometric sensors (Wilson, 2000). The Psychomotor Vigilance Task (PVT) measures correlated positively with heart rate variability while they correlated negatively with heart rate and Automated Neuropsychological Assessment Metrics (ANAM) performance. The results suggested that as participants experience more fatigue, they react more slowly. SA appeared to be affected by fatigue based upon self-reporting over the 36 hour period. Comparisons of the objective SA results were not as clear; however, participants were novice on the SA task and likely experienced learning effects. It has been shown that learning effects exist for even the simplest tasks during long duration studies (Van Dongen, & Dinges, 2000). In comparisons to the other measures, SA had only moderately consistent results. The ANAM Sleepy and PVT Sleepy correlations were all negative, as would be expected; increases in the participant's self-reported sleepiness led to decreases in SA scores. This was most consistently pronounced for the Overall SA. The pilot study demonstrated that the testing protocol supports assessment of Situation Awareness (SA) of fatigued individuals.

Based on the pilot study, a follow-on study was conducted that further investigated the effects of fatigue on SA. Again, participants of the study will experience 36 hours of continued wakefulness in order to simulate chronic fatigue which many sonar operators develop as a result of current shift rotations. To overcome limitations in collecting SA and minimize learning effects, Fleet operators who were fully trained in sonar operation were tested. Physiological data of heart rate, heart rate variability, and eye movements were collected. To assess cognitive performance, the Cambridge Neuropsychological Testing Automated Battery (CANTAB) tasks will be used in place of the PVT and ANAM tests. Workload was compared between two sonar systems by using the subjective NASA Task Load Index (NASA TLX) survey. Having too much work to do can lead to stress and fatigue while having too little work often leads to loss of focus and concentration (Nofi, 2000).

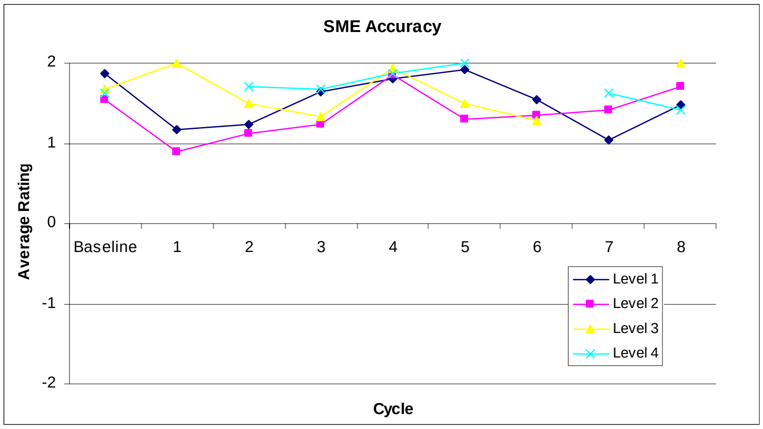

FIGURE 2. SA Accuracy

To measure SA, an SME asked SA probes in an operationally realistic manner during performance of the sonar task and scored the accuracy of their answers. The average for all participants was calculated for each of the four levels of SA (Figure 2). As expected, the participants average accuracy ratings during Cycle 5 for probe level 1 were highest. It is also during this cycle that participants recorded more blinks and saccades per minute. This could indicate that questions of higher SA level could be more affected by hours awake. The number of probes in each cycle ranged from 4 to 10 for Levels 1 and 2 and up to 4 for Levels 3 and 4; there were no Level 4 probes in Cycles 1 and 6, and no Level 3 probes in Cycle 7. The scores in cycles with fewer probes at a level may have been skewed by the limited data.

Participants' scores on Level 1 probes actually improved through time from 1.17 in Cycle 1 to a high of 1.91 in Cycle 5. However, at that point the scores dropped to 1.05 in Cycle 7, before finishing at 1.48 in Cycle 8. Level 2 probes mirrored Level 1, except the high was in Cycle 4, with 1.85, followed by a quick drop back to 1.30. Level 2 received the only accuracy score that was below 1, a 0.89 in Cycle 1. The Level 3 probes received two perfect averages; both Cycle 1 and Cycle 8 averaged a 2.00, with two probes in each cycle. Level 3 lows of 1.33 (Cycle 3) and 1.28 (Cycle 6) were sandwiched around a 1.94 in Cycle 4. Level 4 probes were steady around 1.7 until a 1.88 in Cycle 4 and a 2.00 in Cycle 5, with two probes in each of those cycles. The final two cycles for Level 4 dropped to a low of 1.42.

There appeared to be a peak in performance during Cycle 4 across all levels. It is interesting to note that Level 2 averaged lower than Level 3, which were presumably questions that required more cognitive ability. Level 4 averaged at or above the other levels, except in Cycles Baseline and 8 and the two cycles without Level 4 questions. However, the limited number of Level 3 and 4 probes may have limited the movement of the scores.

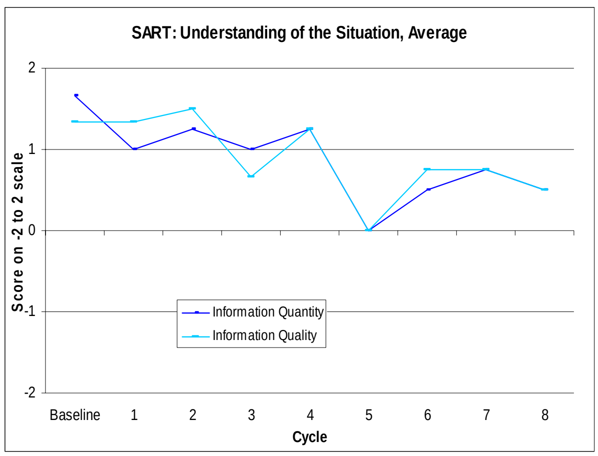

FIGURE 3. Average Scores for Situation Awareness Rating Technique

Participants also responded to questions in the SART survey regarding their perceived demand on attentional resources, supply of attentional resources, and how well they understood the situation (seen in Figure 3). Although there was no noticeable difference in concentration and division of resources over time, participants responded that their arousal and spare capacity decreased as the time they were awake increased. However, there was a slight increase in subjective rating of these measures at Cycle 6. This cycle corresponds to the cycle that received slightly higher ratings of accuracy during the SA tasks and the cycle in which participants tended to make fewer errors when performing the CANTAB tests. The ratings of understanding the situation tended to decrease over time with a slight increase at Cycle 6.

Following each cycle of sonar tasks and SA probes, participants responded subjectively to NASA-TLX workload questions (Average scores of all participants are shown in Figure 4). Measures of effort and frustration tended to steadily increase as the

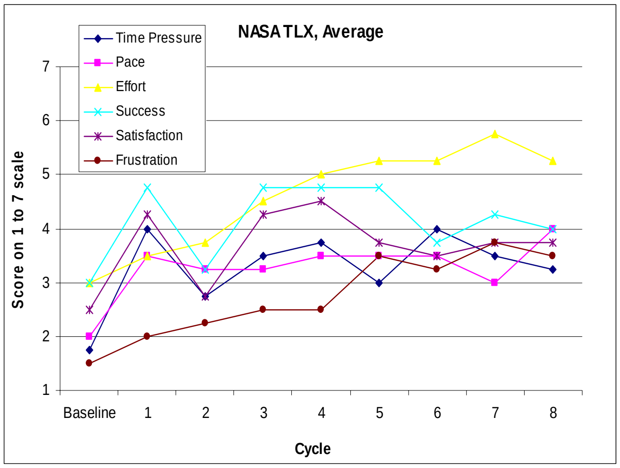

FIGURE 4. Average NASA TLX

participants' time awake increased. The subjective measure of success was high during Cycles 3-5, which correspond to high accuracy ratings, especially for SA Levels 1 and 4.

Before and after completing the sonar tasks, participants were asked to rate their sleepiness level using the Stanford Sleepiness Scale. (See Figure 5.) On average, participants felt that they were the sleepiest both before and after Cycle 5 testing. Participants also tended to rate themselves as being sleepier after each test cycle then they were before starting the testing.

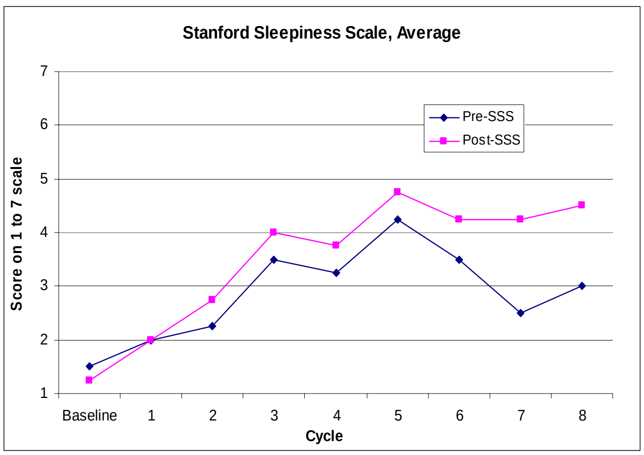

FIGURE 5. Average Scores for Stanford Sleepiness Scale

As hypothesized, task performance degraded over time due to increased fatigue. Several significant relationships were found which support our hypotheses on the relationship between fatigue and situation awareness. Significant positive correlations between the eye tracker data, performance on the CANTAB tests, and subjective measures of effort, frustration, and sleepiness indicate that as the person became more tired, the amount of effort and level of frustration increased. More eye movements were also displayed. Significant negative correlations were found between eye tracker data and participants' perceived rating of spare capacity available for other tasks and their perceived degree of usefulness of the information. Significant positive correlations between the Stanford Sleepiness Scales and subjective measures of effort and frustration suggest that as participants became sleepier, their effort and frustration with the sonar SA task increased. Additionally, significant negative correlations between the Stanford Sleepiness Scales and subjective measures of situation awareness (arousal, spare capacity, information quantity and quality) suggest that as participants became sleepier their supply of attentional resources and their understanding of the situation decreased.

These studies continued to mature the JHU/APL theory on situation awareness, develop a capability/tool to enhance human performance assessment in the areas of workload and situation awareness in real-time, and demonstrated a relationship between fatigue and effective SA. They have demonstrated that it is possible that single summary of Overall SA could be sufficient as a measure of SA, and start to show a relationship between physiological measures and SA. It was found that SA is affected by fatigue, and is highly individualized. This makes it particularly challenging to develop prediction models of how performance will degrade and highlights the need to tailor predictions to specific individuals. It is foundational to developing tools to evaluate potential countermeasures that minimize the effect of fatigue and other stressors in a C2 environment, objectively measuring cognitive state and performance, and providing C2 commanders with a real-time fatigue management system.

Future Activities

The series of studies in cognitive state and physiological sensors continues. JHU/APL is leveraging significant advances in augmented cognition and internal investments in C2 infrastructure to continue assessing warfighter contribution to mission performance in realistic environments providing a framework and scientifically sound evaluation methodology to C2 design and development inclusive of human performance considerations.

The next series of studies will use a physiologic (augmented cognition) sensor suite and investigate the relationships between the physiological measures provided by the sensors and a warfighter's performance under various conditions of stress and fatigue. In the first year, three tasks will be executed: 1) installation of the physiological sensor suite in an existing C2 laboratory environment (the Precision Engagement Transformation Center (PETC), 2) training a small team of people on the use of the sensor suite, and 3) demonstrating the sensor suite with functions from an operational C2 center (like the Maritime Headquarters Maritime Operations Center (MHQ-MOC)). In the second year, the development of objective human performance measurement techniques using a physiological sensor (augmented cognition) suite to uncover the relationships between warfighter cognitive state, physiological measures, and warfighter performance will be undertaken. In addition, integrating these known relationships into the human performance analysis systems, like JHU/APL's CAOC (Combined Air and Space Operations Center) Performance Assessment System (CPAS) will be investigated. The integration of CPAS's process assessment with warfighters' cognitive state would provide a powerful tool for assessing the performance, in real-time, of both training and actual operations.

Development of the augmented cognition measurement capability will start with the installation and integration of the physiological sensors into the PETC. The equipment includes an electrocardiogram, an electroencephalogram, a mobile eye tracker, a galvanic skin response sensor, and software to integrate the sensor outputs to assess cognitive state (provided by Lockheed Martin). Demonstration of the capability to obtain objective human performance measures using physiological devices will be designed and executed in FY07 in a C2 scenario. This will lay the ground work for future experimentation aimed at correlating physiological measures with cognitive effects and performance in the C2 domain. For example, the correlation between physiological measures and the amount of workload a person is handling or their level of situation awareness. This could lead to objective measures of two attributes that cannot usually be measured objectively and non-intrusively. The follow-on experimentation (in year 2 and beyond) will test the software in a controlled environment and in an exercise environment (e.g., C2 Cross Enterprise Initiative (CEI), MHQ-MOC, Tomahawk Joint Fires). Follow-on work will also look at the synergies between the physiological measurements and a process assessment system such as CPAS for a more complete evaluation of C2 activities for use during training and real-time operations, as well as for the design and development of new C2 systems.

REFERENCES

Luu, P., Campbell, G.E. (2005). *"Oops, I did it again": Using Neurophysiological Indicators to Distinguish Slips from Mistakes in Simulation-Based Training Systems.* Paper presented at HCI International 11th Conference on Human-Computer Interaction, 1st International Conference on Augmented Cognition. July 22-27, 2005. Las Vegas, Nevada.

McKneely, J.A, Bevan, M., Cropper, K., Iny, M., & Vaughan, F. (2005). Initial Investigation on Fatigue in Command and Control Situation Awareness: Physiology and Cognitive Performance. Paper presented at The 2006 Command Control Research Technology Symposium. June 20-22, 2006. San Diego, CA.

Nofi, A. A. (2000). Defining and Measuring Shared Situational Awareness. Center for Naval Analyses. CRM D0002895.A1/Final.

Van Dongen, H.P.A., Dinges, D. F., (2000) Circadian Rhythms in Fatigue, Alertness and Performance. In: M. H. Kryger, T. Roth &W. C. Dement (Eds.), Principles and Practice of Sleep Medicine (3rd ed.), pp. 391–399. W. B. Saunders, Philadelphia, Pennsylvania.