Experimental Design and Testing of a Multimodal Cognitive Overload Classifier

Leonard J. Trejo, Neil J. McDonald, Robert Matthews

Quantum Applied Science and Research
5764 Pacific Center Blvd., #107
San Diego, CA 92121
{ltrejo, neil, robm}@quasarusa.com

Brendan Z. Allison
The Scripps Research Institute
10550 North Torrey Pines Road
La Jolla, CA 92037
bci2k2@yahoo.com

Abstract

We report the results of an experiment designed to construct and test a robust multimodal system for automatic classification of cognitive overload. With the assistance of The Scripps Research Institute, twenty-two experienced gamers performed a first-person shooter combat simulation while multiple biosignals and performance were recorded. Signals included EEG, EOG, EMG, ECG, accuracy, and reaction time measures. A kernel partial least squares or KPLS classifier was trained to distinguish subtle differences in EEG spectra within subjects as they pertained to passive viewing, low-difficulty, and high-difficulty simulations. The KPLS classifier was supported and enhanced by algorithms for preprocessing and normalization of EEG and other biosignals. Results indicated that for some subjects, robust classifiers discriminated passive viewing from active performance with accuracies in the range of 99% to 100% and that such models were stable over two test days, using 35-70 min. of training data from Day 1 and testing on data from Day 2. In addition, for some subjects, classifiers discriminated high-difficulty simulations from low-difficulty and passive simulations with stable accuracies of 80% or better across two test days, using 35-70 min. of training data from Day 1 and testing on data from Day 2. We also tested the effect of alcohol intoxication on simulation performance and classifier accuracy. Performance was slightly altered by drinking alcohol to a blood alcohol level of 0.06%, producing more aggressive behavior than with a placebo across subjects. The KPLS classifiers showed remarkable resilience to alcohol effects, considerably less than the effects of day of testing. Not all subjects had such impressive results. To address the lack of generality across subjects, we are currently performing a modified study design that includes provisions for three calibration tasks. These calibration tasks are intended to serve as brief, simple tasks that a user could easily perform at any time as needed to recalibrate algorithms that relate physiology to cognitive state. The tasks include a) eyes-open and eyes-closed for general EEG calibration, b) a mental arithmetic task (divide by twos or sevens) for cognitive load classification, and c) passive viewing of the combat simulation task, for calibration of engagement/disengagement of mental resources. We aim to use the data from the calibration tasks to adapt classification algorithms for variations over time and for inclusion of multiple sensor and data modalities, such as electrocardiographic, electrooculographic, and electromyographic sensor data.

1 INTRODUCTION

The increasing complexity and automation of tasks in aerospace and military operations has also increased the need to test, monitor, and maximize human capabilities, both to efficiently use human resources and to enhance personnel performance. In operational settings, human performance is fundamentally limited by human information processing capabilities. The information processing performance of military personnel is also limited further by the complications of combat or sustained operations, which include, stress, physical strain, sleep deprivation, reduced alertness, mental or physical fatigue, and a wide range of distracting stimuli. To meet the increased demands of today’s complex systems, DARPA launched the Augmented Cognition (AugCog) program, which aims to use advanced technology to maximize human cognitive capabilities. AugCog is a multidisciplinary approach, which includes research on sensors, biosignal analysis, physiological and behavioral models, algorithms for cognitive state estimation and dynamic computational systems for augmenting cognition and performance.
To estimate human cognitive states, some AugCog systems will use physiological signals such as electrocardiograms (ECG), electrooculograms (EOG), electromyograms (EMG) and electroencephalograms (EEG). These biosignals have proven to be useful for real-time detection of hazardous states, such as fatigue, inattention, stress, and cognitive overload and for triggering appropriate countermeasures (St. John, Kobus, & Morrison, 2003; Russell et al., 2005; Wilson et al., 2003). The basic elements of these systems include measures of task stimuli and demands, human performance, and physiological function. These elements are then combined by computer programs that use pattern recognition or signal processing algorithms to model performance or functional states in terms of task or physiological measures. However, until now, the systems that have been developed suffer from three critical limitations, which have prevented operational deployments:

1. The system is not reliable from day to day, and must be re-trained or recalibrated too often or for unacceptably long periods of time. This may result from nonspecific or uncontrolled changes in behavior or physiology that alter the mapping between physiology and performance. Factors such as uncontrolled learning or changes of strategy, uncontrolled variations due to circadian or ultradian rhythms, training data that are too limited to allow for generalization, and finally, mislabeling of training examples arising from inadequate definition of the “ground truth.” This may also result from reliance on a narrow set of measurements, such as exclusive use of a single measure, such as EEG, fNIR, or heart-rate variability, or a set of measures that is too small to capture the variance relevant to the task and human performance.

2. The system is not reliable when the user’s physiology is altered by uncontrolled changes in the external environment. Examples of this include changes in temperature, altitude, motion, lighting, auditory background, or social stress.

3. The system is not reliable when the user’s physiology is altered by uncontrolled changes in the internal environment. Examples of this include consumption of food, voluntary intoxication (drugs, alcohol, and tobacco), involuntary intoxication (chemical, biological, and Radiologic exposure), various illnesses, and trauma.

In this paper we report on a set of controlled experiments in which increasing degrees of control were applied to the three factors listed above. Our aim was to determine the reliability of a system for detecting cognitive overload of expert gamers performing a first-person shooter simulation. The main effects we tested were the effects of test time (day of testing) and intoxication (alcohol consumption) on the accuracy of an algorithm trained to classify successive two-second long segments of task performance as being from states of low workload, moderate to high workload, and overload. We also strove to control a range of other factors and designed the experiments to maximize reliability of the algorithms. In particular, we applied eight types of controls:

1. Control for task learning, by using highly experienced expert participants,
2. Control for time of day and time on task, by keeping these factors constant,
3. Control for test day by requiring participants to perform all task conditions on two or more days.
4. Control of food consumption, by feeding the subjects standardized meals,
5. Control of the test environment, by keeping lighting and temperature constant,
6. Control of inadequate measurement of physiological variance by using a wide set of physiological measures, including EEG, ERP, EOG, EMG, and ECG,
7. Control of inadequate definition of ground truth workload levels by:
   a. having a task expert analyze and score replayed videos of task performance to define momentary task difficulty, workload, and performance,
   b. logging all user inputs, including mouse, keyboard, and a foot pedal,
   c. measuring reaction times to unpredictable probe stimuli (auditory tone bursts), which require the participant to report running sums of kills, deaths, and probe tone bursts.
8. Controlling intoxication directly by requiring participants to perform the same task before and after imbibing a standard dose of alcohol.
In addition, we used a powerful, yet flexible and robust set of algorithms to map physiology and task variables to functional states or performance (Rosipal, Trejo, & Matthews, 2003; Trejo et al., 2006; Wallerius et al., 2005).

2 METHOD

We planned a series of experiments in which we established increasing control over task, performance, and physiological factors (Figure 1). The first two experiments primarily served to evaluate the major effects of workload on performance and physiological variables and to compare two candidate tasks for the third experiment. In Experiment 1, participants performed a first person shooter simulation (Ghost Recon) at four difficulty levels: passive (observe only), easy (4 enemies), medium (10 enemies), hard (20 enemies). Auditory probe tones occurred unpredictably with an average interval of 6 s. There were two attend conditions, ignore or count tones, which were crossed with the four difficulty levels. In the attend tones condition, participants were instructed to silently tally the numbers of kills, deaths, and tones, and report the sums at the end of each trial. The tone-report task served to provide an additional measure of the workload imposed by the game difficulty manipulation. A spectral analysis of EEG activity was performed for each condition. Single-stimulus P3 ERP amplitude analyses were performed for the tones. As compared to the oddball P3, the single stimulus P3 minimizes a secondary task that may distract subjects. A have shown that the single stimulus approach also elicits effective P300s without unnecessary 'standard' tones (Polich & Heine, 1996). Twenty participants performed the tasks. Of these the first six were for determination of measurements and procedures. The remaining 14 participants completed all conditions. On each of two separate days the participants performed the game in three sessions: 1) no drink, 2) drink, 3) no drink, in which the drink was either a placebo (Cool-aid) or a dose of alcohol (Cool-aid with diet 7-Up) estimated to produce a blood level alcohol level of 0.06 %.

In Experiment 2 participants performed a different first person shooter simulation (Battlefield II) at three difficulty levels: passive (observe only), easy (1 enemy), difficult (5 enemies). The auditory tone-report tasks was used again and the two levels (attend, ignore) were crossed with the two difficulty levels. However, in the attend tones condition, participants were instructed to silently tally the numbers of kills, deaths, and tones, press a foot pedal to acknowledge the tones as quickly as possible, and report the tallies after each pedal press. The modified tone-report task served to provide a time-indexed measure of the workload imposed by the game difficulty manipulation. Because the tone-report task was not as time critical as simulated combat, we hypothesized that reaction times to the tones would be greater during high workload periods than during low workload periods. A spectral analysis of EEG activity was performed for each condition. Single-stimulus P3 ERP amplitude analyses were performed for the tones. Six participants performed the tasks. Of these the first two were for determination of measurements and procedures. The remaining four participants completed all conditions. We also asked each participant to complete the NASA TLX workload rating scale after he completed each game condition.

Experiment 3, which is still in progress, replicates the tasks of Experiment 2 and adds new measures that further define the “ground truth” workload levels. For example, we video taped the game and the participant, and then an expert gamer reviewed the tapes and logged ratings of game difficulty, observed workload, and damage taken in 2-s increments. Other analysts created filters that parsed the keystrokes, mouse clicks, shots fired, overall audio level, and computed summary performance and workload metrics using principal factor analyses (e.g. Trejo et al., 1995). To provide a basis for testing the stability of physiological models from day to day, we required each participant to perform the task on four separate days. On each day the participants performed the game in three sessions: 1) no drink, 2) drink, 3) no drink, in which the drink was either a placebo (Cool-aid) or a dose of alcohol (Cool-aid with vodka) estimated to produce a blood level alcohol level of 0.06 %. For calibration task 1 we asked participants to perform three “calibration” tasks, which included repeated 1-minute runs of baseline EEG, EOG, ECG, and EMG, with 30 s of eyes closed and 30 s of eyes open in each run. We also asked participants to perform a mental arithmetic task before, between, and after each game session. The game sessions include a baseline level (passive observation of a recorded game), a low workload level (one enemy), and a high workload level (five enemies). Because results in Experiments 1 and 2 showed that the tone-report task was sensitive to game workload, we kept it in Experiment 3, but eliminated the ignore tones condition. The NASA TLX scale is also being used in Experiment 3. Four participants completed all phases of Experiment 3.

In all tasks, the EEG was recorded form a standard monopolar 10-20 system montage with a digital EEG system (Sensorium) using electrode caps (ECI) with tin electrodes, referred to average mastoids. The EOG was recorded using the same system with taped-on electrodes above and below one eye (vertical, or VEOG) and on the temples.
lateral to each eye (horizontal, or HEOG). The EMG from scalp muscles (masseter, temporalis) was estimated by band-pass filtering the EEG channels at electrodes T7 and T8. We observed that when participants counted tones out loud there was elevated power in the spectrum from 20 Hz to 50 Hz at these sites. In Experiment 3 we also added a bipolar EMG montage over the right trapezius muscle, which has been reported to reflect mental workload in the pattern of activations and gaps between activations of that muscle (Leyman et al., 2001; McLean & Urquhart, 2002). We followed a protocol that had been approved by the Institutional Review Board of The Scripps Research Institute (Wearable Electrophysiologic Sensor Suite for Detection of Neurotoxic Effects, 09/09/2005). All participants signed IRB-approved informed consent forms, and were pre- and post-briefed on the relevant details of the study.

Figure 1. System to develop and test classifiers of mental workload using multimodal physiological responses. Each participant was a highly experienced gamer who had practiced a video game (Ghost Recon, in Expt. 1, Battlefield 2 in Expts. 2-3) to stable levels of performance. In all experiments we recorded, processed, and stored 19-channel EEG, two-channel EOG, and task performance measures. We synchronized and time stamped all measures in 2-s increments. In Experiment 2 we added a task for the participant to react to a command tone using a foot pedal and report game performance statistics and we added a QUASAR capacitive sensor belt to record one channel of ECG. Experiments 1-2 confirmed that game conditions produced controlled levels of workload ranging from 1=disengaged, 2=low, 3=high, and 4=overload. Experiment 3 (in progress) adds new measures that further define the “ground truth” workload levels. For example, we video taped the game and the participant, and then an expert gamer reviewed the tapes and rated game difficulty, observed workload, and damage (gunshots to the player) taken in 2-s increments. In addition we created filters that parsed keystrokes, mouse clicks, shots fired, audio level, and summarized performance and workload metrics using principal factor analyses.
3 RESULTS

3.1 Experiment 1

3.1.1 Performance Effects

We analyzed performance and physiological measures for the effects of game difficulty, tone counting, and alcohol. We hypothesized that game performance would decrease in response to difficulty, tone-counting and alcohol. We also hypothesized the P300 for the tones would decrease, whereas EEG power in the theta band at Fz and alpha band at Pz would increase with difficulty and decrease with alcohol. Analyses of variance (Figure 2) in the 14 participants who complete Experiment 1 indicated the following effects (p<.05). Game performance, measured as the ratio of kills to wounds (damage) decreased as game difficulty increased (F(1,13)=10.27, p<.0005). There was also a significant main effect of tone counting (F(1,13)=7.06, p<.02); however, the effect was the opposite of what we hypothesized. Game performance was consistently lower in the ignore tones conditions than in the count tones conditions. We also tested the mean error rate for the tone counting task, and observed that mean errors also increased significantly with game difficulty. However, neither the main effect of difficulty on tone counting error rates nor the counting × difficulty interaction was significant.

![Figure 2. Effects of game difficulty on game performance and tone-task counting accuracy.](image)

3.1.2 P300

We analyzed P300 as a function of game difficulty, hypothesizing that when subjects attended to the tones, P300 would decrease with game difficulty, whereas when they ignored the tones, P300 would only reflect engagement of resources, thus being high for the passive viewing condition and low for the active conditions. Analyses of variance in the 14 participants indicated that for the attend tones condition, P300 amplitude decreased as game difficulty increased. In the ignore tones conditions, P300 was greater in the passive viewing condition and nearly equal in all the active conditions (Figure 3).

![Figure 3. Effects of electrode ignore/attend tones, and game difficulty on P300 amplitude.](image)
3.1.3 EEG Spectrum

We analyzed the EEG spectrum as a function of game difficulty and hypothesized that alpha rhythm power would decrease as a function of game difficulty whereas theta rhythm power may increase with game difficulty in response to cognitive load (Gevins et al., 1998; Peterson et al., 2006). We found that Alpha 1 and Alpha 2 power were elevated significantly only the passive viewing condition (Figure 4). All other conditions had no significant effect on EEG band power.

![Relative Mean Band Power](image)

Figure 4. Effect of game difficulty on relative mean spectral power in the bands: Delta (1-4 Hz), Theta (4-8 Hz), Alpha 1 (8-10 Hz), and Alpha 2 (10-12) Hz as functions of game difficulty. Numbers of enemies for each condition curve are noted as 0E, 4E, 10E, and 20E. The only significant effect was relatively lower alpha band power in the active conditions (4E-20E) than in the passive (0E) condition.

3.1.4 EEG-Based Classification of Workload

Overall, the performance, ERP, and EEG measures indicated that the game difficulty manipulation produced signs of increasing workload as the number of enemies was increased from zero to 20. This allowed us to substitute game difficulty for workload in choosing random subsets of data to build a classifier for each subject. The approach followed very closely the method described by Wallerius et al. (2005) and will not be repeated in detail here. Briefly EEG power spectra were computed with Welch’s method (1-30 Hz, FFT bin width 0.5 Hz, 256 pt Hanning window, no overlap) using consecutive 2-s EEG segments for each channel. EOG artifacts were removed computationally (Trejo et al., 2006a, 2006b). EEG power was represented in units of dB. The EEG spectra served as input features for a linear KPLS classifier trained to classify one experimental condition versus each other condition. Eight class labels were derived from the four difficulty levels and two tone conditions. Seventy percent of the EEG epochs were used for a training set and 30% were held for testing. Optimum KPLS model order was determined using a Monte Carlo method of cross validation within the training set (100 randomized runs, 90% training set, 10% validation set). Typically less than five PLS components were optimal, and we enforced a limit of 10 components. The criterion for optimality was maximization of the non-parametric signal detection statistic, $A'$, which is a relatively bias-free estimate of classification sensitivity (Grier, 1971). The multiple 1 vs. 1 models were combined with a nonlinear voting scheme models were highly successful in using EEG spectra to classify workload levels in individual participants, averaging from $A' = .79$ to $A' = 0.93$. Results of the PLS models were tabulated in terms of $A'$ for the 14 participants (Table 1). In general, the closer together two workload levels in terms of enemies or counting tones, the lower the discrimination of the classes by the KPLS models.
3.1.5 Observations

We observed several shortcomings of Ghost Recon and other design factors in Experiment 1. First the Ghost Recon game settings cannot easily be modified. Second, Ghost Recon must be paused and reset frequently. Third, Ghost Recon cannot record good-quality videos of game-play or performance. Fourth, the gamer exit interviews indicated that the workload imposed by Ghost Recon even with 20 enemies did not often reach overload levels. For these reasons we investigated other games and selected Battlefield 2 to replace GR. Battlefield 2 did not suffer from the many of the shortcomings noted for Ghost Recon and it allows us to drive workload to high levels in a more controllable manner than is possible in Ghost Recon.

Table 1. Average Linear KPLS Classification Accuracy of Ghost Recon 2 Workload Levels Using 20-Channel EEG Spectra From 2-s Epochs. The dependent variable is $A'$, a measure of sensitivity to class differences, which ranges from 0 to 1.0, where 1.0 is perfect discrimination. The data represent classifier discrimination performance between the levels indicated by rows and columns.

<table>
<thead>
<tr>
<th></th>
<th>Ignore tones</th>
<th>Count tones</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4 no count</td>
<td>10 no count</td>
</tr>
<tr>
<td>4 enemies</td>
<td>0.82</td>
<td>0.88</td>
</tr>
<tr>
<td>10 enemies</td>
<td>0.82</td>
<td>0.88</td>
</tr>
<tr>
<td>20 enemies</td>
<td>0.87</td>
<td>0.85</td>
</tr>
</tbody>
</table>

3.2 Experiment 2

3.2.1 Performance Effects

Experiments 2 and 3 were designed to test repeated measures in single participants, so statistics of the means in various effects were not analyzed. Instead, our focus was on optimizing and streamlining the workload manipulations and testing the day-to-day validity of the KPLS models, with and without the complicating effects of alcohol. In all participants, we observed that increasing workload on the Battlefield 2 game was associated with decreased foot pedal accuracy and increased pedal reaction time. Interestingly there were no substantial overall trends in numbers of kills or deaths as difficulty increased. Finally, alcohol intoxication did not appear to impair overall game performance. In fact alcohol may have led to more aggressive behavior, as the number of total game points and kills was greater in alcohol conditions than placebo conditions.

3.2.2 P300

We analyzed P300 as a function of game difficulty, under the same hypotheses as in Experiment 2. Analyses of variance in the 14 participants indicated that for the ignore tones condition; P300 amplitude was larger in the passive condition than in the 1 or 5-enemy conditions. Surprisingly, no significant differences in P300 amplitude were observed in the attend tones condition. P300 amplitudes were generally lower after alcohol consumption than in the placebo conditions.

3.2.3 EEG and KPLS Models of Workload

Analyses of the multivariate EEG/ECG/EOG/EMG for Experiment 2 are still in progress. Our initial results indicate two trends. First KPLS models of EEG and ECG that classify workloads decrease in accuracy from the first test day to a subsequent test day. The loss of accuracy ranges from about 10% to 40%, resulting in models that are 60% to 80% accurate. These models are trained exclusively with data from one day and tested on another day. When models are tested with hold-out data from the same test day, accuracy is higher as it was for Experiment 1. Second, the
addition of a second modality, ECG, may improve model generalization to the second day of testing. We tried combining EEG and ECG features in a single input matrix and using KPLS to relate this input to the class labels. We also constructed separate KPLS classifiers using five time domain measures and 15 frequency domain measures of heart rate and heart rate variability. We used custom software scripts (Igor PRO) to compute measures equivalent to those computed by a commercial/academic Windows software package (University of Kuopio, 2006). It is too soon to provide definitive answers, but our initial analyses show that redundant EEG and ECG classifiers track different aspects of workload in the Battlefield 2 game. In particular, we found a stabilizing effect of combining the classifiers in one participant, which kept accuracy at about 80% on both test days. Additional stabilization was observed when we used a short segment of the passive viewing condition data from Day 2 to calibrate the model from Day 1 by adding these observations to the input matrix of features. This manipulation led to a similar stabilizing effect (Table 2).

The classification was binary for 1-enemy (low workload) or 5-enemy (high workload) conditions. Models were trained and tested on the same day (Models A, B, F) or different days (Models D, E, G). For models E and G a portion (3.5 min) of passive viewing condition data from Day 2 were used to calibrate the model for Day 1. The effect of alcohol on the classifier is reflected in the same-day tests for Day 2 (Models F and G). In general testing on different days led to decreases in model accuracy, or $A'$. Comparing models A and D shows a decrease in accuracy ranging from .12 to .17. However adding a small portion of calibration data from Day 2 (Model E) restores accuracy to within about .02 to .04 of the same-day test data. Alcohol has no effect on model accuracy in this participant as shown by comparing models F and G, which remained in the high range of accuracy (.85 to .88) regardless of alcohol exposure.

### 3.3 Experiment 3

Experiment 3 is in progress. This experiment will test several hypotheses, including the concept of accurate momentary estimates of “ground truth” workload, multimodal classification, classifier calibration and stabilization. The overall design is detailed in Figure 1. Note that a wide range of behavioral, task, and physiological measures will be estimated and used to develop optimal class labels and feature sets for multimodal classification.

| Table 2. Linear KPLS Classification Accuracy of Battlefield 2 Intra-/Inter-Session Workload Levels Using 18-Channel EEG Spectra From 2-s Epochs (S206) (channels T7 and T8 were omitted due to EMG contamination) |
|---|---|---|---|---|---|
| **Model** | **Training Set** | **Test Set** | **Training Set** | **Test Set** |
| **Low (0 + 1 Enemy)** | **High (5 Enemies)** | **Low (0 + 1 Enemy)** | **High (5 Enemies)** |
| A | 1, 2, 3, 7, 8 | 5, 6, 10, 11 | 1, 2, 3, 7, 8 | 5, 6, 10, 11 |
| B | 1, 2, 3 | 5, 6 | 7, 8 | 10, 11 |
| D | 1, 2, 3, 7, 8 | 5, 6, 10, 11 | Day 2: 1, 2, 3 | Day 2: 5, 6 |
| E | 1, 2, 3, 7, 8 + Day 2: 1, 2 | 5, 6, 10, 11 | Day 2: 1, 2, 3 | Day 2: 5, 6 |
| F | Day 2: 1, 2, 3, 7, 8 | Day 2: 5, 6, 10, 11 | Day 2: 1, 2, 3, 7, 8 | Day 2: 5, 6, 10, 11 |
| G | Day 2: 1, 2, 3 | Day 2: 5, 6 | Day 2: 7, 8 | Day 2: 10, 11 |

### 4 DISCUSSION

Using physiology to track operator functional states is a challenging proposition. While some states, such as fatigue (Wallerius et al., 2005; Trejo et al., 2006a) or drowsiness are relatively easy to detect, higher cognitive states defined as “workload”, “attention”, or “effort” have proven more difficult to estimate from their physiological correlates.
We sought to address this challenge in a series of experiments in which we imposed increasing levels of control over factors that may cause models to fail. We also progressively increased the set of physiological, behavioral, and environmental measures. We draw two distinct conclusions:

1. Video game analogs of complex tasks such as military first-person combat can induce significant and controllable variance in workload and performance in expert users. Physiological correlates of this variance can discriminate different workloads with high accuracy under certain conditions. These conditions include single-user models and same-day training and testing of the models.

2. Variation in model accuracy from day to day is unacceptably large, ranging up to as much as 40% losses in accuracy. However, two stabilization methods appear to provide some increases in model test-retest reliability: a) the use of small segments of calibration data from the second testing day, the inclusion of multiple classifiers trained to recognize workload-related patterns in different modalities, such as EEG, ECG, EOG, and EGM measures. We are now working on a range of stabilization methods, including additional modalities of EOG and ECG as well as new calibration methods.

The experiments described here are studies on the development of multimodal classification. They are enabling us to have precise control over tasks and measures in the model development stage. The complex design shown in Figure 1 is our present development system. This system allows us to generate a wide range of controlled experiments and test multiple measures and algorithms for estimating operator functional states. However, our ultimate aim is to design an automated system that minimizes hardware and software. We have already designed and built several of the key hardware components. The finished product will include a small body-worn processor, an array of unobtrusive sensors, and a wireless transmitter. A remote receiver, which may be a small portable computer or a software agent in a wireless command and control network, will interpret the sensor data and provide advisory or cautionary signals. These signals will gauge the functional states of individual operators and may be used for interventions or adjustments. Another design goal is the requirement that an individual user can be fitted” for the sensors and algorithms in one session and perform a brief “warm up” task to recalibrate the system at the beginning of each work shift. We envision such systems as products that will increase the efficiency of human-system operations and increase safety by detecting hazardous states that are not directly observable, such as excessive mental workload.

5 ACKNOWLEDGEMENT

We gratefully acknowledge the support of the Army Research Office SBIR Contract, “Wearable Electrophysiologic Sensor Suite for Detection of Neurotoxic Effects, Phase II,” and the useful comments of its technical monitor, Dr. Elmar T. Schmeisser.

6 REFERENCES


