

# EEG-Based Estimation and Classification of Mental Fatigue

Leonard J. Trejo<sup>1\*</sup>, Karla Kubitz<sup>2</sup>, Roman Rosipal<sup>1,3</sup>, Rebekah L. Kochavi<sup>4</sup>,  
Leslie D. Montgomery<sup>5</sup>

<sup>1</sup>Pacific Development and Technology, LLC, Palo Alto, USA

<sup>2</sup>Towson University, Towson, USA

<sup>3</sup>Department of Theoretical Methods, Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovak Republic

<sup>4</sup>Stinger Ghaffarian Technologies, Inc., Moffett Field, USA

<sup>5</sup>LDM Associates, San Jose, USA

Email: [\\*ltrejo@pacdel.com](mailto:ltrejo@pacdel.com)

Received 28 January 2015; accepted 3 April 2015; published 9 April 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Mental fatigue was associated with increased power in frontal theta ( $\theta$ ) and parietal alpha ( $\alpha$ ) EEG rhythms. A statistical classifier can use these effects to model EEG-fatigue relationships accurately. Participants ( $n = 22$ ) solved math problems on a computer until either they felt exhausted or 3 h had elapsed. Pre- and post-task mood scales showed that fatigue increased and energy decreased. Mean response times rose from 6.7 s to 7.9 s but accuracy did not change significantly. Mean power spectral densities or PSDs of  $\theta$  and  $\alpha$  bands rose by 29% and 44%, respectively. A kernel partial least squares classifier trained to classify PSD coefficients (1 - 18 Hz) of single 13-s EEG segments from alert or fatigued task periods was 91% to 100% accurate. For EEG segments from other task periods, the classifier outputs tracked the time course of the development of mental fatigue. By this measure, most subjects became substantially fatigued after 60 min of task performance. However, the trajectories of individual classifier outputs showed that EEG signs of developing fatigue were present in all subjects after 15 - 30 minutes of task performance. The results show that EEG can track the development of mental fatigue over time with accurate updates on a time scale as short as 13 seconds. In addition, the results agree with the notion that growing mental fatigue produces a shift away from executive and attention networks to default mode and is accompanied by a shift in alpha frequency to the lower alpha band.

## Keywords

EEG, Mental Fatigue, Alertness, Drowsiness

---

\*Corresponding author.

## 1. Introduction

The purpose of this study was two-fold: a) to identify the features of the electroencephalogram or EEG that change in an orderly and reliable manner with the development of mental fatigue; and b) to construct a practical system that can use EEG features to accurately estimate the instantaneous degree of mental fatigue over a performance period of up to three hours. Our approach consisted of five key steps. First, we adopted a clinical-theoretical framework that allowed us to distinguish *mental* fatigue from general fatigue, sleepiness or other factors that influence task performance. Second, we adopted a task and a set of procedures for inducing mental fatigue in the participants of the study. Third, we adopted psychometric tests that can independently validate the ability of the task to induce mental fatigue. Fourth, we selected an EEG measurement system that is practical enough to use in operational settings and sensitive enough to detect changes in EEG features with the development of fatigue. Finally, we adapted an algorithm for estimation of mental fatigue from EEG features that is robust in the face of signal corruption from noise or sensor loss and accurate enough to classify states of mental fatigue in individual human subjects with a few seconds of EEG recordings.

There are countless occupations today in which fatigued individuals routinely operate complex, high-risk systems. Such systems may be as common as driving a car or operating a forklift, or as rare as controlling traffic at a busy airport or docking two spacecraft in orbit. This undesirable state of affairs has contributed to many well publicized—and some not so well publicized—disasters. For example, operator fatigue was involved in the Three-Mile Island and Chernobyl nuclear power plant disasters, the EXXON Valdez grounding, and the Guantanamo Bay airline crash (U.S. Congress, Office of Technology Assessment, 1991; Hall, 2000; Dinges, 1995). “There is [also] much evidence that [operator] fatigue has contributed to serious incidents and accidents in industrial operations ... and to accidents in *all* [emphasis added] modes of transportation” (Dinges, 1995: p. 8). Analyses of crash data confirm that fatigue, drowsiness and inattention pose the greatest known risks to automobile driver and passenger safety, surpassing all other known risks including alcohol use and secondary tasks such as cell phone usage (Dingus et al., 2006; Sagberg et al., 2004). Accordingly, there continues to be much scientific interest in assessing, monitoring, and predicting operator fatigue (Gevins et al., 1995; Kennedy, 1953; Wilson & Fisher, 1991; 1995; Russo, Stetz, & Thomas, 2005).

Unfortunately, there is no standard definition of fatigue. Clinically, fatigue is best defined as difficulty in performing voluntary activities (Chaudhuri & Behan, 2004). The main psychological symptoms are subjective feelings of low energy and motivation, which may be accompanied by physical weakness. Drowsiness aside, fatigue is still a broad concept, which includes mental, motor, metabolic, endocrine, sensory, and other physiological or environmental components (Chaudhuri & Behan, 2004; Winningham, 1996). We prefer the point of view that fatigue is distinct from drowsiness, or the feeling of a need to sleep, although fatigue and drowsiness are often felt at the same time (Hurd, 2007). Hockey (2013) considered that mental fatigue arises not so much from excessive work and effort as from a lack of motivation to continue tasks that become less and less rewarding over time. In this view, fatigue is healthy in the sense that it can help prevent motivational fixation on current activities. For example, the onset of fatigue can force a break in attention to a monotonous and unrewarding task, and allow a person to weigh the benefit of continuing that task or redirecting behavior towards activities that may be more rewarding. Nevertheless, our society often requires people to engage in long-duration tasks with low immediate reward value, such as long-haul truck driving, piloting aircraft or various monitoring activities. A better understanding of the neurophysiology of alertness-fatigue transitions may aid in task design and analysis. Ultimately, knowing precisely what changes in an individual’s EEG during such transitions could be used with biofeedback to train operators of fatigue-inducing systems to self-regulate these transitions and avoid accidents (e.g., Beatty et al., 1974).

In the present study, we will attempt to experimentally isolate *mental fatigue* from general fatigue because complex, high-risk tasks such as piloting aircraft and vehicles impose minimal physical demands and heavy cognitive demands on human operators. In such tasks, mental fatigue jeopardizes performance and safety through a variety of cognitive impairments, including: a) decreased vigilance (Mast & Heimstra, 1964; Šípoš, 1971; Smit, Eling, & Coenen, 2004); b) decreased secondary task performance (van der Hulst, Meijman, & Rothengatter, 2001); c) impaired problem solving (Horne, 1988; van der Linden, Frese, & Meijman, 2003; van der Linden, Frese, & Sonnentag, 2003); d) reduced situation awareness (Vidulich et al., 1994); and e) decreased tolerance of risk and effort (Barth, Holding, & Stamford, 1976; Brown, Tickner, & Simmonds, 1970; Neri, Shappell, & De John, 1992; Shingledecker & Holding, 1974).

As with all mental states, clinicians and neuroscientists cannot sharply define mental fatigue in terms of specific physiological processes. So mental fatigue is usually defined as a subjective assessment that is associated with various behavioral and physiological symptoms. For example, we can define mental fatigue as the unwillingness of alert, motivated subjects to continue performing a specific mental task (Montgomery, Montgomery, & Guisado, 1995). Then at times before, during, or after performance of the task, we can validate this definition with psychometric tests.

Although mental fatigue is currently assessed subjectively, there is a pressing practical need for objective measures to estimate, monitor, and predict it. Changes in the EEG clearly signal the onset and stages of sleep, but more recent studies have shown that EEG patterns can also signal the onset of mental fatigue in fully awake persons. For example, a separate analysis of data from this study found evidence for three statistically separable stages of mental fatigue development: an alert state, a “normal” state (neither alert nor fatigued), and a mental fatigue state (Trejo et al., 2007). However, despite many studies to date, the EEG signs of fatigue are still not sharply defined and using EEG for operational fatigue estimation is still a formidable challenge.

This state of affairs motivated us to rigorously test the accuracy and robustness of EEG-based estimation of mental fatigue. We used a two-pronged approach: first we assessed the common across-subjects pattern of changes in power spectral density or *PSD* of multichannel EEG over extended periods of mental arithmetic performance. Second, we tested the feasibility of using the PSDs of EEG to classify mental fatigue states at a sub-minute time scale in individual subjects. More specifically, we trained, cross-validated, and tested statistical learning algorithms for classifying brief segments of EEG recordings as pertaining to self-reported states of alertness or mental fatigue.

To assess EEG changes common to all subjects, we tested the hypothesis that the PSDs of theta, alpha, and other EEG bands would co-vary regularly with the degree of mental fatigue that subjects experienced. While many studies have considered EEG correlates of general fatigue, drowsiness and sleepiness, our study focused on inducing mental fatigue through sustained performance of a cognitive task for up to three hours, while minimizing other effects. For example a study of EEG classification of mental fatigue using support vector machines from subjects who performed an auditory vigilance task reported classification accuracies ranging from 87.2% to 91.2% (Shen et al., 2008). However, these results may have confounded sleepiness and mental fatigue, as the subjects had endured 25 hours of sleep deprivation before performing the vigilance task. Other studies lead us to expect significant fatigue-related changes in PSDs of the frontal midline theta band and the parietal alpha band (e.g., Lorist et al., 2005; Gevins et al., 1998; Barwick et al., 2012). So we specified null hypotheses that EEG power in specific theta and alpha bands would remain constant over the course of a fatigue-inducing task.

To test the feasibility of using EEG to monitor mental fatigue in individuals, we trained a supervised machine learning algorithm to extract information-rich features from brief segments of EEG recordings and classify them as pertaining to states of alertness or mental fatigue. The algorithm is known as *kernel partial least squares* or *KPLS*, which we have applied successfully to classification of a wide range of data types including linear and nonlinear classification of EEG and event-related potentials (Rosipal & Trejo, 2001; Rosipal et al., 2001; Wallerius et al., 2005; Trejo, Matthews, & Rosipal, 2006). The features we extract using KPLS are a set of orthogonal basis vectors, which exhibit high covariance with states of alertness or fatigue. Except for this covariance constraint the basis vectors are computed using methods similar to those used in principal components analysis or factor analysis. We find the necessary and sufficient set of basis vectors using algorithms (de Jong, 1993; Rosipal, Trejo, & Matthews, 2003) that extract each vector sequentially, in descending rank of its covariance with an output measure, which in this case is a distinction between alertness and fatigue. Each extracted basis vector is a unique linear combination of multichannel EEG power spectral densities. We can then iterate this feature extraction procedure with cross validation methods to optimize the number of extracted components by striking a robust balance of classification accuracy and generality. We apply the KPLS algorithm in this study to accurately and robustly estimate mental fatigue in the framework of developing a useful system for continuous monitoring of mental fatigue under working conditions.

## 2. Methods

### 2.1. Experimental Design

We tested several hypotheses about the association of subjective moods, observed behavior, performance, and physiological measures during continuous performance of mental arithmetic for up to three hours. We manipu-

lated a single factor, that is, time on task, and used a repeated measures design. Subjective moods were indexed by the Activation Deactivation Adjective Checklist (AD-ACL; Thayer, 1986; 1989) and the Visual Analogue Mood Scales (VAMS; Stern, 1997) questionnaires. Observed behavior included ratings of activity and alertness from videotaped recordings of each participant's performance. The performance measures were response time and response accuracy, recorded in real time by the task computer. The physiological measures were derived from spontaneous EEGs, including: a) theta activity (4 - 8 Hz) at Fz (both average power and peak amplitude in the theta band) and b) alpha activity (8 - 13 Hz) at Pz (both average power and peak amplitude in the alpha band).

## 2.2. Participants

We collected data from 33 individuals from the NASA Ames Research Center community. All participants gave their informed consent using a form approved by the NASA Ames Research Center Institutional Review Board and were paid for their participation. We report only the data from the 22 participants who met the following inclusion criteria: they had to have a) been able to wear an EEG electrode cap without significant discomfort; b) stayed awake for at least 90 minutes; c) provided stable EEG recordings within a tolerance of  $\pm 100 \mu\text{V}$ . We dismissed participants who failed to meet these criteria. Six of the 22 eligible participants were excluded from analyses. One was excluded because he violated experimental protocol and wore a watch during the testing, and attended to the watch repeatedly. Five more were excluded because their response times were extremely slow (and consequently they provided too few EEG epochs for analysis). The remaining 16 participants included 12 males and 4 females with a mean age of 26.9 (SD = 7.4) years. Also, according to self-reports, all of the participants had normal vision and hearing. Fourteen of the 16 participants reported being right-handed and two reported being left-handed.

## 2.3. Mental Arithmetic Task

Participants sat in front of a computer with the right hand resting on a 4-button keypad (Neuroscan STIM pad, Compumedics USA, El Paso, TX). Arithmetic summation problems, consisting of four randomly generated single digits, three operators, and a target sum (e.g.,  $4 + 7 - 5 + 2 = 8$ ), were displayed on a computer monitor (cathode ray tube) continuously until the subject responded (Figure 1). Only addition and subtraction were used, and equations for which answers were obvious (such as those including several repeated digits) were excluded. The participants: a) solved the problems; b) decided whether their calculated sums were less than, equal to, or greater than the target sums provided; c) indicated their decisions by pressing the appropriate key on the keypad. The keypad buttons were labeled “<”, “=”, and “>”, respectively. Subjects were instructed to answer as quickly as possible without sacrificing accuracy. After a response, there was a 1 s inter-trial interval, during which the monitor was blank. Participants performed the task until either they quit from exhaustion or three hours had elapsed. As can be seen in Table 1, all participants completed at least 6 of the 12 15-min blocks and 11 participants completed all 12 of the 15-min blocks (i.e., the total 3-hour period).

## 2.4. Activation Deactivation Adjective Checklist

Thayer's AD-ACL (Thayer, 1986; 1989) is a multi-dimensional checklist reflecting perceptions of activation.

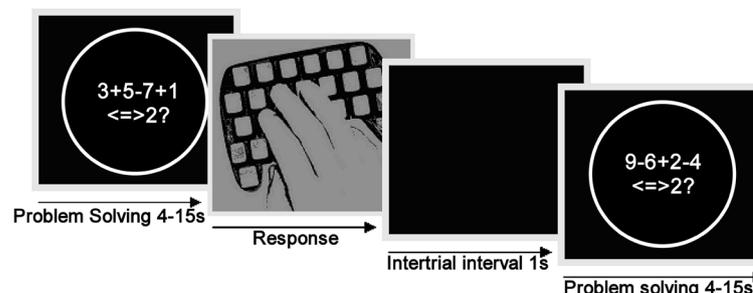


Figure 1. Schematic diagram of events in the mental arithmetic task.

**Table 1.** Number of participants per 15-min block.

Block	1	2	3	4	5	6	7	8	9	10	11	12
Time (min)	15 - 30	30 - 45	45 - 60	60 - 75	75 - 90	90 - 105	105 - 120	120 - 135	135 - 150	150 - 165	165 - 180	180 - 205
Participants	16	16	16	16	16	16	15	14	14	13	12	11

Individuals respond to 20 items using a 4-point rating scale (definitely feel, feel slightly, cannot decide, and definitely do not feel). The scoring procedure includes four subscales: energy (reflects general activation), tiredness (reflects general deactivation), tension (reflects high preparatory arousal), and calmness (reflects low preparatory arousal). The AD-ACL is a reliable and valid subjective method (Thayer, 1986; 1989).

## 2.5. Visual Analogue Mood Scales

The VAMS (Stern, 1997) measure eight specific mood states, including afraid, confused, sad, angry, energetic, tired, happy, and tense. The VAMS have a neutral schematic. That is, they have a “mood-neutral” face (and word) at the top of a 100 mm vertical line and they have a “mood-specific” face (and word) at the bottom of the line. Individuals mark the point along the line that best illustrates how they feel at present. Scores range from 0 to 100, with 100 indicating the maximum level of the mood and 0 indicating the minimum level of a mood. Like the AD-ACL, the VAMS are also reliable and valid (Nyenhuis et al., 1997; Stern, 1997).

## 2.6. Observed Activity and Alertness

Activity and alertness were measured by visual inspection of videotapes of each participant’s performance. The videotapes showed combined overall scene and facial views of the participants. For each 15-min interval, a single rater judged levels of alertness and activity (unnecessary motion) on a five-point scale (Table 2). The alertness scale considered the subjective appearance of sleep, sleepiness, dozing off, distraction, yawning, and general alertness. The activity scale considered the frequency of subject motion, including moving in the chair, fidgeting, tapping, or shaking. The rater was blind to the specific aims of the study, although not to the time interval being rated (as the videotaped segments were viewed sequentially). However, changes in activity and alertness were related to obvious changes in behavior and, as a result, were fairly easy to quantify. Moreover, a second rater randomly rated a subset of the videotapes and found good agreement with the ratings. Ratings of activity and alertness were tested for correlations with response time, accuracy, and EEG spectral measures.

## 2.7. EEG Activity

EEG was recorded continuously using 32 Ag/AgCl electrodes embedded in an elastic fabric cap (Quik-Cap™, Compumedics USA, El Paso, TX). The electrode cap was placed on the participant according to the manufacturer’s instructions. The reference electrodes were digitally averaged mastoids and the ground electrode was located at AFz. Vertical and horizontal electrooculograms (VEOG and HEOG) were recorded using bipolar pairs of 10 mm Ag/AgCl electrodes (one pair superior and inferior to the left eye; another pair to the right and to the left of the orbital fossi). Impedances were maintained at less than 5 k $\Omega$  for EEG electrodes and 10 k $\Omega$  for EOG electrodes. The EEG was amplified and digitized with a 64-channel Neuroscan Synamps™ system (Compumedics USA, El Paso, TX), with a gain of 1000, sampling rate of 500 s<sup>-1</sup> and a pass band of .1 to 100 Hz. Amplifiers were calibrated with a 50  $\mu$ V signal prior to each testing session. The signals were digitized and stored on hard disk drives by a computer equipped with Neuroscan Scan 4.2 software (Compumedics USA, El Paso, TX) and archived on optical media (CD-R).

## 2.8. Procedures

Participants: a) were given an orientation to the study; b) read and signed an informed consent document; c) completed a brief demographic questionnaire (age, handedness, hours of sleep, etc.); d) practiced the mental arithmetic task for 10 minutes; and e) were prepared for data collection by having the electrode cap, EOG, and reference electrodes applied. They then completed the pretest self-report measures (i.e., the AD-ACL and VAMS) and performed the mental arithmetic task until either three hours had elapsed or volitional exhaustion

**Table 2.** Scale used to quantify observed activity and alertness.

Value	Activity	Alertness
1	Little or no movement (0% - 10% time)	Asleep
2	Occasional movement (10% - 25% time)	Dozing off but still responding
3	Intermittent movement (25% - 50% time)	Distracted, yawning or only somewhat alert
4	Frequent movement (50% - 75% time)	Completely awake and mostly alert
5	Constant movement (75% - 100% time)	Completely awake and fully alert

had occurred. Task termination was followed by the completion of post-test self-report measures and participant debriefing.

## 2.9. Data Processing

The EEGs, initially processed using Neuroscan Scan 4.2 Edit™ (Compumedics USA) software, were: a) submitted to an algorithm for the detection and elimination of eye-movement artifact; b) visually examined and blocks of data containing artifact were manually rejected; c) epoched around the stimulus (i.e., from -5 s pre-stimulus to +8 s post-stimulus); d) low pass filtered (50 Hz; zero phase shift; 12 dB/octave roll off); and e) submitted to an automated artifact rejection procedure (i.e., absolute voltages > 100  $\mu$ V). The overall single-epoch rejection rate was 47%. The cleaned and filtered epochs resampled at 128 Hz using a non-aliasing method. EOG artifact was removed by using wavelet-denoised VEOG and HEOG signals as predictors of the artifact voltages at each EEG electrode in a multivariate linear regression. The ocular artifact rejection algorithm consisted of four main steps: a) wavelet smoothing of the EOG channels; b) blink detection; c) computing the linear regression of the EEG on the EOG channels; and d) subtracting the resulting weighted EOG from the EEG channels. These steps were repeated for each single EEG epoch. The wavelet smoothing process decomposed the EOG channels by using a Daubechies (DWT-8) wavelet filter. The lowest 2% of the coefficients were retained and the signal was reconstructed from the retained coefficients (e.g., Trejo & Shensa, 1999). The voltage of each sample in the VEOG signal was then compared to an absolute voltage threshold set at 100  $\mu$ V and marked as a blink or non-blink accordingly. A multivariate linear regression algorithm was used to calculate the regression weights for the EEG on the VEOG and HEOG channels. Separate regression weights were estimated for blink and non-blink epochs. The final step was a linear subtraction of the weighted VEOG and HEOG signals from each EEG channel. EEG PSDs were estimated with Welch's periodogram method at 833 frequencies from 0 - 64 Hz. Peak and average power in the theta and alpha bands were measured at electrodes Fz and Pz, respectively.

## 2.10. Classification Procedures

We classified single EEG epochs using kernel partial least squares decomposition of multichannel EEG spectra coupled with a discrete-output linear regression classifier, or *KPLS-DLR*. Through extensive side-by-side testing of EEG data, we found that *KPLS-DLR* is just as accurate as *KPLS-SVC*, which uses a support vector classifier for the classification step (Rosipal, Trejo, & Matthews, 2003). *KPLS* selects the reduced set of orthogonal basis vectors or *components* in the space of the input variables (EEG spectra) that maximizes covariance with the experimental conditions. *DLR* finds the linear hyperplane in the space of *KPLS* components that separates the classes. In a pilot study (Montgomery, Montgomery, & Guisado, 1995) and in the present study, we found that the first 15 minutes on task did not produce substantial mental fatigue, whereas mental fatigue was substantial in the final 15 minutes. So we randomly split EEG epochs from the first and last 15-min periods into equal-sized training and testing partitions for classifier estimation. Only the training partition was used to build the final models. The number of *KPLS* components in the final models was set by five-fold cross-validation. The criterion for *KPLS-DLR* model selection was the minimum classification error rate summed over all (five) cross-validation subsets.

## 2.11. Statistical Analyses

The data were analyzed using either singly or, when appropriate, doubly multivariate repeated measures analyses of variance with time of measurement as a within-subjects factor. Doubly multivariate analyses were carried out

for AD-ACL subscale scores (energy, tension, calmness, and tiredness), VAMS subscale scores (afraid, confused, sad, angry, energetic, tired, happy, and tense), behavioral observation data (observed activity and alertness), theta activity measures (peak and band-average amplitudes), and alpha activity measures (peak and band-average amplitudes). Singly multivariate analyses of variance were carried out for response time and accuracy. For the doubly multivariate analyses, significant multivariate F-ratios were decomposed using single degree of freedom, within-subjects contrasts. For the singly multivariate analyses, Huynh-Feldt-corrected degrees of freedom and  $p$ -values were reported (i.e., because of sphericity). In both analyses, partial  $\eta^2$  values were reported as effect size estimators.

### 3. Results

#### 3.1. Self-Report Analyses

The AD-ACL subscale scores were analyzed in a doubly multivariate ANOVA with time of measurement (i.e., pretest vs. posttest) as a within-subjects factor. The main effect of time of measurement was significant ( $F(4, 5) = 10.4, p < .01, \eta^2 = .89$ ). Within-subjects contrasts showed significant linear trends for energy ( $F(1, 8) = 6.46, p < .04, \eta^2 = .45$ ), calmness ( $F(1, 8) = 21.3, p < .002, \eta^2 = .73$ ), and tiredness ( $F(1, 8) = 6.38, p < .04, \eta^2 = .44$ ). Energy decreased from a pretest mean of 12.0 (standard deviation (SD) = 4.1) to a posttest mean of 8.6 (SD = 3.7). Calmness decreased from a pretest mean of 16.8 (SD = 1.6) to a posttest mean of 14.1 (SD = 1.8). Tiredness increased from a pretest mean of 10.1 (SD = 4.3) to a posttest mean of 15.3 (SD = 5.7). There was also a non-significant linear trend for tension ( $F(1, 8) = .92, p < .37$ ). Thus, the AD-ACL data indicate that our manipulation decreased general activation (i.e., self-reported energy) and preparatory arousal (i.e., self-reported calmness) and increased general deactivation (i.e., self-reported tiredness).

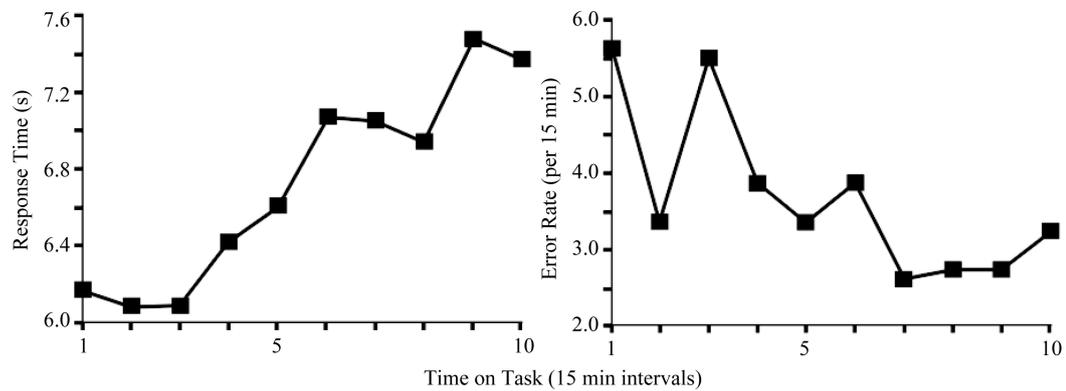
The VAMS subscale scores (i.e., for afraid, confused, sad, angry, energetic, tired, happy, and tense) were analyzed in a doubly multivariate ANOVA with time of measurement (i.e., pretest vs. posttest) as a within-subjects factor. The main effect of time of measurement was non-significant (multivariate  $F(8, 1) = 1.31, p = .59$ ). This analysis suggests that our manipulation, despite its effects on activation and arousal, did not influence other, more durable, moods.

#### 3.2. Behavior Analyses

The behavioral observations (i.e., observed activity and alertness) were analyzed in a doubly multivariate ANOVA with time of measurement (i.e., 15-min periods) as a within-subjects factor. The main effect of time of measurement was significant ( $F(18, 178) = 3.70, p < .0005, \eta^2 = .27$ ). This analysis suggests that time on task influenced behavior (i.e., observed activity and alertness levels). Moreover, time on task had a progressive effect on behavior. Within-subjects contrasts showed a linear decrease in alertness ( $F(1, 10) = 10.4, p < .009, \eta^2 = .51$ ) and a linear increase in activity ( $F(1, 10) = 5.88, p < .04, \eta^2 = .51$ ). Alertness decreased from a mean of 5.00 (SD = .00) in the first 15-min period to a mean of 2.43 (SD = .98) in the last 15-min period. Activity increased from a mean of 1.36 (SD = .51) to a mean of 2.45 (SD = 1.30), respectively.

The response times (RT) were analyzed in an ANOVA with time of measurement (i.e., 15-min periods) as a within-subjects factor. The main effect of time of measurement was significant (Huynh-Feldt corrected  $F(3, 39) = 3.78, p < .03, \eta^2 = .24$ ). This analysis suggests that time on task influenced performance. Moreover, time on task had a progressive effect on performance (Figure 2). Within-subjects contrasts showed a significant linear increase in RT ( $F(1, 12) = 8.29, p < .01, \eta^2 = .41$ ) rising from a mean of 6.70 s (SD = 2.18) in the first 15-min period to a mean of 7.87 s (SD = 2.64) in the last 15-min period. We found the same pattern of significant effects for RT analyzed in an ANOVA with fraction of artifact-free trials (i.e., 1<sup>st</sup> 100, middle 100, and last 100) as a within-subjects factor.

Response accuracy was analyzed in an ANOVA with time of measurement (i.e., 15-min periods) as a within-subjects factor. The main effect of time of measurement was not significant (Huynh-Feldt corrected  $F(5, 43) = 1.74, p = .14$ ). Response accuracy was also analyzed in an ANOVA with fraction of artifact-free trials (i.e., 1<sup>st</sup> 100, middle 100, and last 100) as a within-subjects factor. The main effect of number of trials was not significant (Huynh-Feldt corrected  $F(2, 19) = 2.84, p = .09$ ). This analysis suggests that, despite its effects on other aspects of behavior, time on task did not have a substantial influence on response accuracy.



**Figure 2.** Effects of time on task on response time and accuracy. Response times trended linearly upwards over time beginning after Block 3, or 45 minutes on task. Error rates declined over time but the trend was not significant.

### 3.3. EEG Analyses

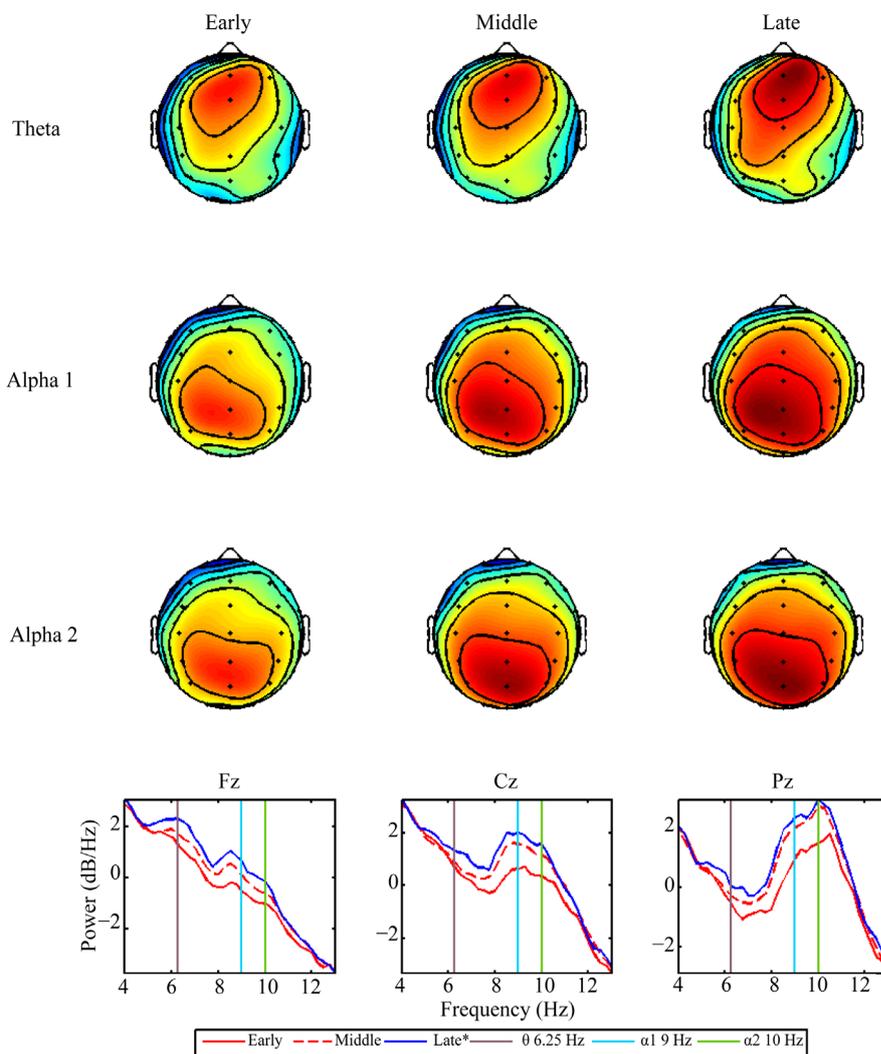
Because the small number of participants precluded analysis of the full montage of electrodes, we examined the average spectra for pronounced effects (**Figure 3**) and focused our analyses on the electrodes that had maximal average voltages for the alpha and theta bands (i.e., Fz for theta and Pz for alpha). Electrode Fz shows a marked increase in theta power spectral density near 6 Hz. Electrode Pz shows marked increases in alpha PSD with peaks near 9 Hz (alpha 1) and 10 Hz (alpha 2). The increase in theta power is steady over the three time periods, whereas the increase in alpha 1 and alpha 2 is large between early and middle time periods and smaller between middle and late time periods. The relative distribution of average alpha 1 power is more anterior than alpha 2, with alpha 1 showing a maximum at Cz and alpha 2 showing a maximum at Pz.

The changes in frontal midline theta (i.e., average PSDs and peak amplitudes at Fz) were analyzed in a doubly multivariate ANOVA with time of measurement (i.e., 15 min-periods) as a within-subjects factor. Each subject's theta peak was measured individually. The main effect of time of measurement was significant (multivariate  $F(18, 178) = 2.05, p < .01, \eta^2 = .17$ ). Average power in the theta band (4 - 8 Hz) increased from a mean of  $199.36 \mu\text{V}^2/\text{Hz}$  ( $\text{SD} = 97.50$ ) in the first 15-min period to a mean of  $256.58 \mu\text{V}^2/\text{Hz}$  ( $\text{SD} = 135.57$ ) in the last 15-min period. Peak amplitude in the theta band increased from a mean of  $272.4 \mu\text{V}^2/\text{Hz}$  ( $\text{SD} = 146.0$ ) in the first 15-min period to a mean of  $390.8 \mu\text{V}^2/\text{Hz}$  ( $\text{SD} = 227.1$ ) in the last 15-min period. This analysis suggests that theta increased with time on task. Moreover, this analysis suggests that time on task had a progressive effect on frontal midline theta activity. Within-subjects contrasts showed significant linear increases in average theta PSDs ( $F(1, 10) = 7.42, p < .01, \eta^2 = .48$ ) and in peak theta amplitudes ( $F(1, 10) = 9.31, p < .01, \eta^2 = .48$ ).

The changes in midline parietal alpha activity (i.e., average PSDs and peak amplitudes at Pz) were analyzed in a doubly multivariate ANOVA with time of measurement (i.e., 15-min periods) as a within-subjects factor. For each subject the peak alpha power was measured individually without respect to distinctions of alpha 1 and alpha 2 bands shown in **Figure 3**. The main effect of time of measurement was significant (multivariate  $F(18, 178) = 2.20, p < .005, \eta^2 = .18$ ). Average power in the alpha band (8 - 13 Hz) increased from a mean of  $307.4 \mu\text{V}^2/\text{Hz}$  ( $\text{SD} = 434.3$ ) in the first 15-min period to a mean of  $459.0 \mu\text{V}^2/\text{Hz}$  ( $\text{SD} = 593.9$ ) in the last 15-min period. This analysis suggests that alpha power increased progressively with time on task. Within-subjects contrasts showed significant linear increases in average alpha PSDs ( $F(1, 10) = 6.07, p < .03, \eta^2 = .38$ ). Peak alpha amplitudes increased and trended similarly, but not significantly so ( $F(1, 10) = 4.11, p = .07$ ).

### 3.4. Classification

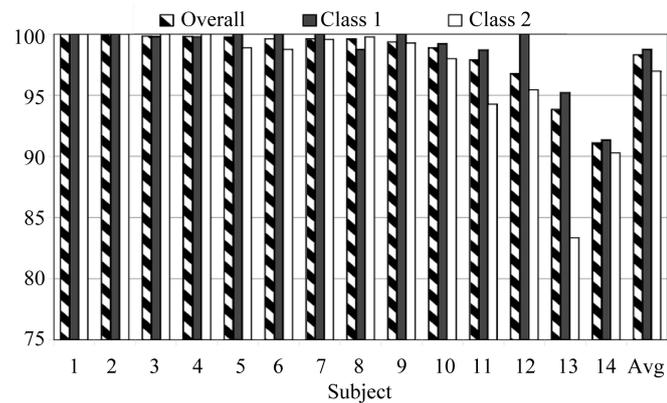
We applied our classification procedure to EEG recordings from 14 of the 16 subjects (two subjects had too few EEG epochs for model estimation). The EEG epochs were synchronized with the onset of each math problem, extending from  $-5$  s to  $+8$  s relative to each stimulus onset. As such there was some overlap among the EEG segments. However a second analysis of 3.5 s segments with no overlap produced highly similar results, so we will focus only on the long-epoch results. We also reduced the likelihood of electromyogram artifacts by low-pass filtering the EEG with an 18-Hz cutoff.



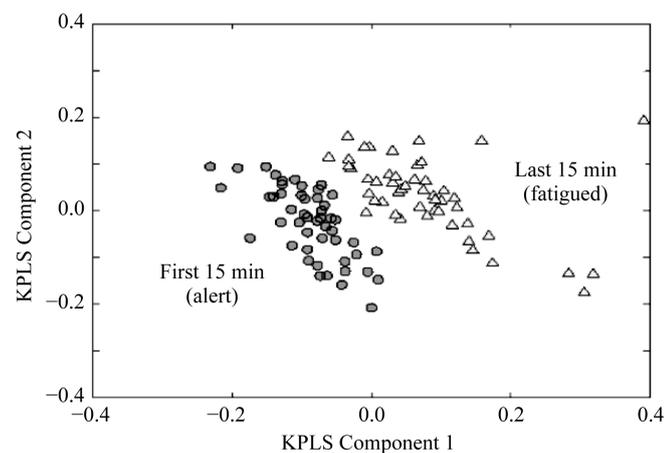
**Figure 3.** Grand average EEG power spectra and topographic maps at peak frequencies in the theta band (peak 6.25 Hz), alpha 1 band (peak 9 Hz) and alpha 2 band (peak 10 Hz) across all subjects as a function of time on task. Time periods for each spectrum or plot were *early* = first 15-min block, *late* = last 15-min block, *middle* = all other blocks. Electrode Fz shows a marked increase in theta power spectral density (PSD) near 6 Hz. Electrode Pz shows a marked increase in alpha PSD with peaks near 9 Hz (alpha 1) and 10 Hz (alpha 2). The color scales of the topographic maps are in units of dB/Hz and range from blue to red with the following limits: theta =  $-3.6$  to  $2.4$ , alpha 1 =  $-4.5$  to  $2.7$ , alpha 2 =  $-5.3$  to  $3.0$ . All topographic maps in this report used a spherical spline interpolation algorithm (Perrin et al., 1989) and the APECSgui toolbox (Pacific Development & Technology, 2012) with an adaptation of the *eev\_topoplot* function from EEGLAB (Delorme & Makeig, 2004).

For each subject we made a KPLS-DLR model using either linear or Gaussian (nonlinear) kernels and selected the best model as described above. On average linear kernels had slightly better results than Gaussian kernels, so we focus on linear kernels here. Classification accuracies (Figure 4) across both classes for 18-Hz filtered EEG ranged from 91.12 to 100% (mean = 98.30, Table 3). The corresponding range for 11-Hz filtered EEG was 89.53 to 98.89% (mean = 98.30%). The number of KPLS components ranged from 1 to 4 (mean 2.77) for 18-Hz EEG and from 1 to 5 (mean 3.76) for 11-Hz EEG (Table 3).

With as few as two components, the separation of classes was usually evident from the distribution of KPLS scores for single EEG epochs. The test-set data for the first and last 15-min blocks occupied distinct regions in the space of the KPLS scores (Figure 5).



**Figure 4.** Classification accuracies for 14 subjects and the averages across subjects with 1 - 18 Hz band pass. Hatched bars show test set accuracies for overall classification; black bars are for Class 1 (alert) epochs, and white bars are for Class 2 (fatigued) epochs.



**Figure 5.** Example of KPLS scores predicted for single-trial EEG spectra for early (dark circles) and late (light triangles) blocks of the mental arithmetic task in one subject (S3).

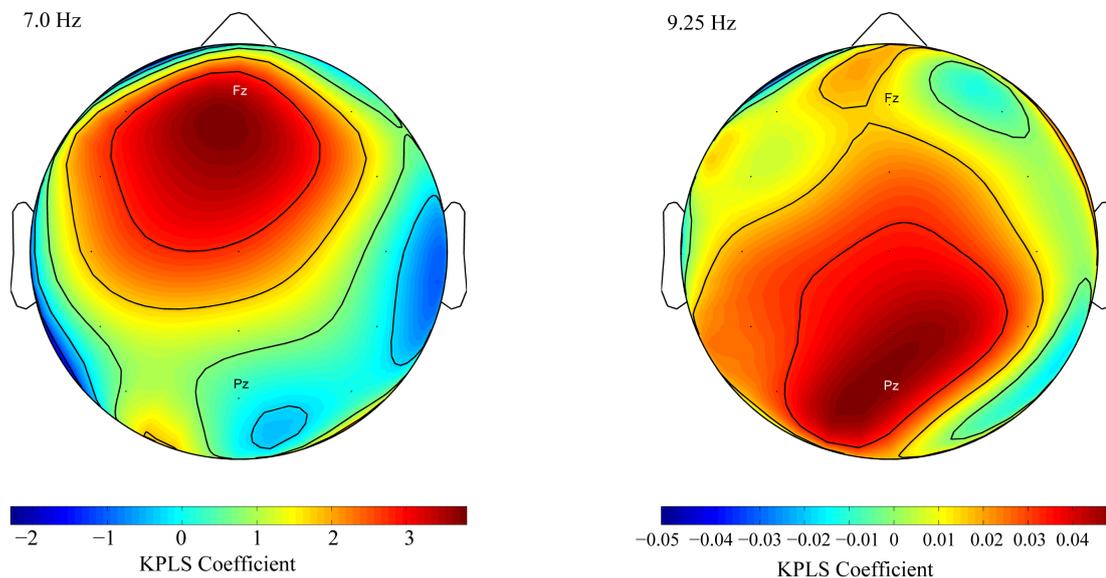
**Table 3.** KPLS-DLR classification accuracies (test proportion correct) by filter cutoff and class membership.

Low pass cutoff	TPC train	TPC test	TPC Class 1	TPC Class 2	Mean number of components
11 Hz	99.96	97.01	97.37	95.57	3.76
18 Hz	99.86	98.30	98.78	96.97	2.77

The scalp topography of the KPLS weights can serve as an indicator of which regions or electrodes strongly influence classification. For example, by plotting the KPLS model coefficients at the theta and alpha band peaks of the EEG spectra in one subject (Figure 6), we see that frontocentral midline sites are important for classification in the theta band, with a maximum near Fz. In the alpha band, the discriminating electrodes were concentrated near Pz.

### 3.5. KPLS Model Prediction

We also examined the predictive validity of the KPLS-DLR models by testing them with data from the first nine intervening 15-min periods (between first and last). The behavior of the classifiers for these periods was consistent with an orderly, progressive migration of single-trial KPLS predictions from the non-fatigued to the fati-



**Figure 6.** Topographical maps of the coefficients for the first KPLS component at the peak frequencies in theta band (left, 7.0 Hz) and alpha band (right, 9.25 Hz) for one subject (S7). The colored areas are spherical spline interpolated values, with the largest values in red and the smallest in blue. Electrodes nearest the maxima were Fz for the theta band and Pz for the alpha band.

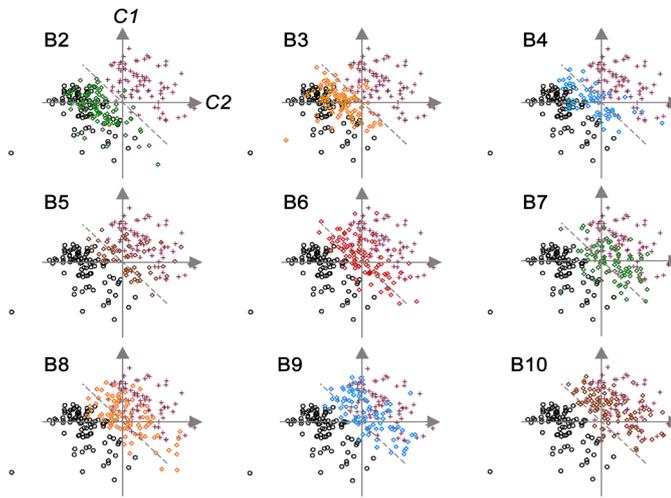
gued class. This observation agrees with the trends we observed in response times, EEG measures, and behavioral observations. We examined these patterns of migration by inspecting graphs of the predicted scores for the first two components of the KPLS models for single subjects (Figure 7). Initially, the predicted points overlapped with the region occupied by the non-fatigued training set. Over time, the predicted points shifted towards the fatigue region. For each of 11 subjects, we computed the center of mass (mean) of the KPLS classification scores for each of 10 fifteen-minute blocks (Figure 8). Negative scores represent alert states and positive scores represent fatigue states. For most subjects there is a progressive shift of the means toward the fatigue state, with some intermittent reversals. Three subjects entered fatigue between 15 - 30 min whereas one subject did not enter fatigue until between 105 - 120 min. Although the majority of subjects (n=6) did not cross the zero line between classified alertness and fatigue until between 60 - 75 min, all subjects show progressive transitions towards the fatigue state in the first 15 - 30 minutes.

## 4. Discussion

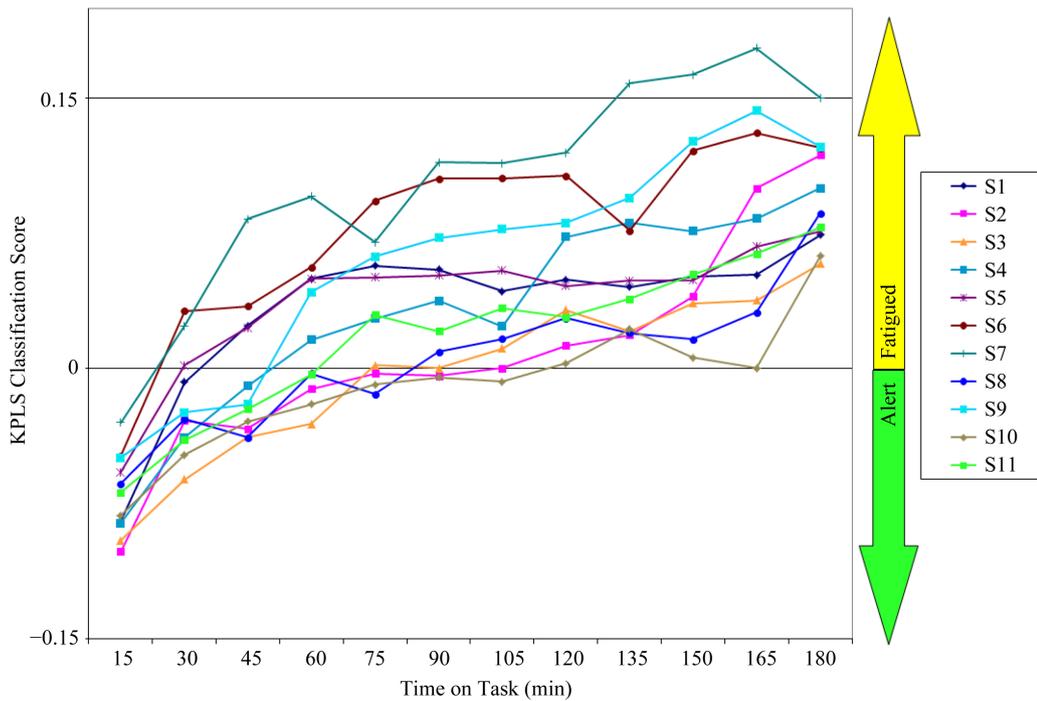
Our results showed that the power spectral densities of theta and alpha EEG bands co-varied with the degree of fatigue that subjects experienced. Power in both theta and alpha bands increased significantly between the beginning and end of the task. These increases were confirmed with behavioral and subjective reports as pertaining to alert and fatigued conditions. Moreover, we found that we could accurately model and estimate fatigue as a function of EEG spectral features using KPLS-DLR, a statistical learning theory based classifier.

### 4.1. Behavioral Measures

Time on task produced decreased general activation (i.e., self-reported energy) and preparatory arousal (i.e., self-reported calmness) and increased general deactivation (i.e., self-reported tiredness) but did not significantly influence other moods. These effects support the assertion that our task produced a state of substantial mental fatigue. Observations of the participants showed that behavioral activity progressively increased while alertness progressively decreased over time. Moreover, there was a progressive slowing of response times. However, time on task did not significantly influence response accuracy. Together, these results suggest that our subjects experienced mental fatigue, but did not sacrifice accuracy as might be expected if motivation had waned. The moderate, general increases in RT over time also indicate increasing mental fatigue, but not a severe increase in RT, as might be expected if lapses or microsleep episodes had occurred frequently.



**Figure 7.** Estimating the development of fatigue over time in one subject (S3). KPLS scores were predicted for EEG epochs from nine 15 minute blocks between the training set blocks (B2 - B10). Block 2 = 15 - 30 min, block 3 = 30 - 45 min, block 4 = 45 - 60 min, ... , block 10 = 135 - 150 min. Black circles and purple crosses are the KPLS C1 and C2 scores of single EEG epochs from fatigued (B1) and non-fatigued (B12) training sets, respectively. Colored diamonds are the KPLS C1 and C2 scores ( $x = C1, y = C2$ ) of single EEG epochs for intervening 15-minute blocks 2 - 10. In this subject, the drift of the orange diamonds in B3 away from the black circles and towards the purple crosses marks the onset of fatigue after 30 - 45 minutes on task. By the tenth block most predicted scores fell in the fatigue region, as defined by the training data set.



**Figure 8.** Means of the KPLS classification scores for each 15-minute block in the 11 subjects who completed 12 blocks or 180 minutes of continuous mental arithmetic performance. For subject S3 the points in the range 30 to 150 min (Blocks 2 - 10) correspond to the centers of the clouds of points shown moving from the alert region to the fatigue region in **Figure 7**. Negative scores represent alert states and positive scores represent fatigue states.

Prolonged vigilance appears to have different effects on error rates than fatigue arising from prolonged cognitive effort such as mental arithmetic. For example, in a simulated sonar task, an EEG-based alertness index followed closely the time course of detection error rates (Makeig & Inlow, 1993). However, we found that cognitively fatigued subjects continued to perform with low and stable error rates for up to three hours. Our task differed from the sonar task by virtue of having strong cognitive demands, whereas the demands of the sonar task were mainly sensory or perceptual (detecting sonar sounds). The main behavioral impact of *cognitive* fatigue in our arithmetic task appeared to be a slowing of mental processes, because response times trended significantly and progressively higher over time. Because the response demands of our task were simple and highly practiced, we do not consider that cognitive fatigue substantially slowed response selection processes. Instead, our results are more consistent with the view that cognitive fatigue slows down central executive functions such as working memory and decision making as compared to perception and response-selection.

## 4.2. EEG Measures

Our EEG analyses indicate that time on task had a progressive influence on PSDs of frontal midline theta and parietal alpha activity. PSDs in both of these EEG bands increased as a function of time on task. Our inspection of the EEG spectra did not indicate effects outside the theta and alpha bands. In particular, there were no indications of effects at 14 Hz or in the beta band. The lack of effects in bands above 14 Hz also discounts electro-myographic artifacts or other broad-band noise sources as determinants of the theta and alpha effects. Progressive increases in alpha and theta band powers were also observed in a study of extended performance of an air traffic control task (Krishnan, Dasari, & Ding, 2014). In general agreement with our findings, the EEG changes in the air traffic control task indicated alertness-to-fatigue “transition times” after about 70 minutes of task performance. In our mental arithmetic task such transitions may occur earlier, in as few as 15 - 30 minutes of continuous performance. Task differences may account for the shorter transition time observed here, or the KPLS-derived EEG measures we used may be more sensitive than those used by Krishnan et al.

Our results do not support an overall slowing of the EEG in mental fatigue, as much as they indicate specific increases in frontal midline theta and midline parietal alpha power. However, the scope of our analyses does not allow us to rule out specific fatigue-related decreases in individual peak alpha frequency or relative proportions of power in lower and upper alpha bands. More recent analyses across several studies of sustained attention tasks suggests that fatigue may induce a shift in the frequency of the alpha rhythm accompanied by a relative shift of the alpha sources towards more anterior locations (Rosipal et al., 2009; Trejo et al., 2010; Rosipal et al., 2013; Zaidel et al., 2013). A similar finding was observed by Barwick et al. (2012) when mental fatigue developed in a Stroop interference task. Studies of individual alpha frequency changes should be a fertile area in future mental fatigue research.

One explanation of the fatigue-induced increases in alpha band power derives from a neurophysiological model about the significance of varying alpha band power levels. The model states that “when patches of neurons display coherent activity in the alpha band, a depressed state of active processing of information in the underlying cortical neuronal populations can be assumed to exist” (Pfurtscheller & Lopes da Silva, 2005: pp. 1008-1009). This model is consistent with the idea that the likelihood of depressed or reduced processing surrounding the onset of a task-relevant stimulus increases as a function of time on task or in relation to subjectively perceived mental fatigue.

Another possible explanation for increasing alpha power concerns the timing of our EEG spectral measures. Our EEG epochs were synchronized with the onset of each stimulus presented in the mental arithmetic task. This means that the EEG features associated with task event processing will tend to dominate the analysis, as compared to EEG features that are randomly selected over time. It is possible that over time on task, alpha oscillations became progressively more synchronized with task events. For example, if bursts of alpha tended to occur randomly early on but became regularly linked to stimuli or responses later. Our present analyses provide no evidence for or against this explanation as we did not look at event-related synchronization measures, but this again seems a fertile area for future analyses.

The hypothesized fatigue-related increase in frontal midline theta band PSDs was supported by our analyses. One current hypothesis states that increases in theta band PSDs are associated with increases in the demands of cognitive processes (Klimesch, 1999; Pfurtscheller & Lopes de Silva, 2005). In our mental arithmetic task we consider that the critical cognitive processes are working memory, long-term memory, and selective attention. A large and diverse body of experimental reports has substantiated that working memory (WM) capacity and speed

of access to long-term memory (LTM) are critical limiting processes in cognitive performance (Anderson, 1995). In particular, a theoretical framework has emerged, which relates alpha, theta, and beta activity in the anterior cingulate cortex and the hippocampus to WM and LTM (Gevins et al., 1997; Klimesch et al., 1997; Mecklinger, Kramer, & Strayer, 1992). Distinctions are now recognized among the spatial and frequency parameters that characterize: spatially diffuse theta rhythms associated with sleepiness, spatially focused, theta rhythms associated with cognitive effort and WM load, lower alpha rhythms (8 - 10 Hz) associated with relaxed wakefulness, and focused higher alpha rhythm (10 - 12 Hz) suppression associated with spatial and non-spatial shifts of attention (reviewed in Klimesch, 1999). These experimental and theoretical results are generally consistent with the theta band effects we observed here.

Another model for the increase in frontal midline theta band power states that the frontal executive network may go “off-line” during states of mental fatigue. In this view, the observed increases in theta power would indicate that rhythmic theta represents an “idling” state and the frontal executive network is in a standby mode. In support of this view, Lorist et al. (2005) found evidence of reduced frontal event-related potential amplitude with time on task in subjects performing a mentally fatiguing switch task for two hours. Another study, in which functional magnetic resonance were images recorded while participants performed a psychomotor vigilance test task for 20 minutes, also found evidence of reduced frontal-parietal activity with time on task (Lim et al., 2010). Activity in the fronto-parietal network was lower during post-task rest compared to pre-task rest, and cerebral blood flow decreased in this network in a manner correlated with a decline in task performance. The view that increasing theta signals reduced activity in the frontal executive network goes hand in hand with the view that increasing alpha activity signals reduced activity in parietal regions. Both theta and alpha rhythms increasing could also suggest a withdrawal from executive and attentive processing to more activity in the default mode network. Modeling of EEG-fMRI correlations at the network level is still a young field that must explain a complex and subtle set of effects (Murta et al., 2015), but some evidence shows that default mode network activation appears to be more highly related to beta-band activity than to alpha or theta (Laufs et al., 2003). However, where correlations with default mode network activation and alpha were obtained, it was with low frequency alpha (Jann et al., 2009; 2010), which would agree with the notion that growing mental fatigue produces a shift away from executive and attention networks to default mode and is accompanied by a shift in alpha frequency to the lower alpha band. Perhaps it is in such flights away from task performance to default mode that a person finds time to re-evaluate goal priorities, as predicted by Hockey’s (2013) motivational model of mental fatigue. We strongly encourage new research to sort out the reasons for the significant effects of mental fatigue on increased alpha and theta band powers, as well as their relationship to functional brain networks.

### 4.3. Classification and Prediction

Overall, our KPLS-DLR classification of single-trial EEG epochs ranged from 90% to 100% with a mean of 97% to 98%. The performance of these classifiers is highly accurate for single trials, and may serve as the basis for predictive models of mental fatigue in operational settings. Inspections of the predictive behavior of the KPLS models showed an orderly relationship between the scores and time on task as well as between the scores and correlated behavioral, subjective, and performance measures. In the 11 subjects that we examined for model predictions, our classifier generalized in an orderly manner to data from the 15-minute blocks in between the initial and final 15-minute blocks used to train the classifier. Specifically, the ensembles of KPLS scores used to estimate mental states from single EEG epochs in these intervening blocks exhibited an orderly progression from alertness to fatigue. According to the estimates of our classifier, subjects fatigued at widely different rates, but most subjects entered the region of KPLS scores corresponding to fatigue after about 60 min on task. Although the focus of the present study was on binary classification of alert vs. fatigue states, we have also assessed the clustering of KPLS-derived features with these data and have found evidence that the development of mental fatigue comprises three distinct states of alertness: alert, normal, and fatigued (Trejo et al., 2007). Future studies should examine whether such states correspond to physiologically distinct network activities or represent levels of activation within a fixed set of networks.

The present study also shows that the individual EEG spectral signatures and patterns of change are useful for estimation of cognitive fatigue. The general trend across subjects in EEG spectral change is explained by an alpha + theta model, with both alpha and theta PSDs being greater in fatigued states than in normal states. However, the estimation of momentary cognitive fatigue states from single EEG trials of 13 s duration was most effective with multi-electrode frequency spectra of individual subjects. The KPLS algorithm correctly identified

features of EEG spectra for each individual, and accurately classified over 90% of single EEG trials as having come from a period of alertness or mental fatigue in most subjects.

Our results suggest several important implications for future EEG studies of mental fatigue. First, as others have found (e.g., Galbraith & Wong, 1993; Smith et al., 2001), EEG classification algorithms benefit greatly by being both individualized and sensitive to multiple sensors and frequency bands. The development of simpler and more general models, which apply to a broad set of subjects, will require considerable additional research. Another well-known problem in the applied EEG community is that the performance of classification algorithms from day to day, or at different times of day is unstable (Trejo, Matthews, & Allison, 2007; Wilson & Russell, 2003a). Additional research is needed to develop methods for stabilizing the link between EEG features and mental states such as fatigue or alertness over long periods of time. Finally, some type of saliency analysis, such as our maps of spectral features (Figure 6), is important in order to determine which fundamental aspects of the EEG yield the most accurate predictions (Wilson & Russell, 2003b). Additional research is also needed to develop quantitative methods for saliency analyses that are meaningful across studies and feature extraction or classification algorithms.

## Acknowledgements

This research was supported in part by the NASA Aerospace Operations Systems Program directed by Dr. Malcolm Cohen, the Advanced Estimation of Cognitive Status Project and the Neurosensory Optimization of Information Transfer Project sponsored by Drs. Elmar T. Schmeisser and Frederick D. Gregory of the Life Sciences Division of the US Army Research Laboratory's Army Research Office and grant number APVV-0668-12 to Dr. Roman Rosipal from the Slovak Research and Development Agency and grant number MZ 2013/46-SAV-6 from the Ministry of Health of the Slovak Republic. Additional support was provided by the Independent Research and Development Fund of Pacific Development and Technology, LLC. We would also like to acknowledge the support of the NASA/ASEE Faculty Fellowship and the NASA Ames Educational Associates programs for supporting Dr. Karla Kubitz. We also acknowledge the contributions of Richard W. Montgomery and Bernadette Luna for theoretical discussions and use of the mental arithmetic task, Bryan L. Matthews for extensive experimental, analytical and programming support, and Melina L. Trejo for support of data analysis tasks. Finally we thank Dr. Maria Ondrejčaková for a thoughtful review of the manuscript and her help in preparing the final version of the paper for submission.

## References

- Anderson, J. R. (1995). *Cognitive Psychology and Its Implications: Fourth Edition*. New York: Freeman.
- Barth, J. L., Holding, D. H., & Stamford, B. A. (1976). Risk versus Effort in the Assessment of Motor Fatigue. *Journal of Motor Behavior*, 8, 189-194. <http://dx.doi.org/10.1080/00222895.1976.10735071>
- Barwick, F., Arnett, P., & Slobounov, S. (2012). EEG Correlates of Fatigue during Administration of a Neuropsychological Test Battery. *Clinical Neurophysiology*, 123, 278-284. <http://dx.doi.org/10.1016/j.clinph.2011.06.027>
- Beatty, J., Greenberg, A., Deibler, W. P., & O'Hanlon, J. F. (1974). Operant Control of Occipital Theta Rhythm Affects Performance In a Radar Monitoring Task. *Science*, 183, 871-873. <http://dx.doi.org/10.1126/science.183.4127.871>
- Brown, I. D., Tickner, A. H., & Simmonds, D. C. V. (1970). Effect of Prolonged Driving on Overtaking Criteria. *Ergonomics*, 13, 239-242. <http://dx.doi.org/10.1080/00140137008931137>
- Chaudhuri, A., & Behan, P. O. (2004). Fatigue in Neurological Disorders. *The Lancet*, 363, 978-988. [http://dx.doi.org/10.1016/S0140-6736\(04\)15794-2](http://dx.doi.org/10.1016/S0140-6736(04)15794-2)
- De Jong, S. (1993). SIMPLS: An Alternative Approach to Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, 18, 251-263. [http://dx.doi.org/10.1016/0169-7439\(93\)85002-X](http://dx.doi.org/10.1016/0169-7439(93)85002-X)
- Delorme, A., & Makeig, S. (2004). EEGLAB: An Open Source Toolbox for Analysis of Single-Trial EEG Dynamics Including Independent Component Analysis. *Journal of Neuroscience Methods*, 134, 9-21. <http://dx.doi.org/10.1016/j.jneumeth.2003.10.009>
- Dinges, D. F. (1995). An Overview of Sleepiness and Accidents. *Journal of Sleep Research*, 4, 4-14. <http://dx.doi.org/10.1111/j.1365-2869.1995.tb00220.x>
- Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., & Knippling, R. R. (2006). *The 100-Car Naturalistic Driving Study, Phase ii—Results of the 100-Car Field Experiment Performed by Virginia Tech Transportation Institute, Blacksburg, VA, Sponsored by National Highway Traffic Safety Administration*, Washington DC, April 2006, DOT HS 810 593 - April 2006.

- Galbraith, G. C., & Wong, E. H. (1993). Moment Analysis of EEG Amplitude Histograms and Spectral Analysis: Relative Classification of Several Behavioral Tasks. *Perceptual and Motor Skills*, 76, 859-866. <http://dx.doi.org/10.2466/pms.1993.76.3.859>
- Gevins, A., Leong, H., Du, R., Smith, M. E., Le, J., DuRousseau, D., & Libove, J. (1995). Towards Measurement of Brain Function in Operational Environments. *Biological Psychology*, 40, 169-186. [http://dx.doi.org/10.1016/0301-0511\(95\)05105-8](http://dx.doi.org/10.1016/0301-0511(95)05105-8)
- Gevins, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S., Du, R., & Rush, G. (1998). Monitoring Working Memory Load during Computer-Based Tasks with EEG Pattern Recognition Methods. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40, 79-91. <http://dx.doi.org/10.1518/001872098779480578>
- Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-Resolution EEG Mapping of Cortical Activation Related to Working Memory: Effects of Task Difficulty, Type of Processing, and Practice. *Cerebral Cortex*, 7, 374-385. <http://dx.doi.org/10.1093/cercor/7.4.374>
- Hall, J. (2000). Opening Remarks. National Transportation Safety Board, Public Hearing on Aircraft Accident, American Airlines Flight 1420, Little Rock, Arkansas, 1 June 1999. Accessed 27 January 2015. <https://app.nts.gov/news/speeches/hall/jhc000126.htm>
- Hockey, R. (2013). *The Psychology of Fatigue*. Cambridge University Press. <http://dx.doi.org/10.1017/cbo9781139015394>
- Horne, J. A. (1988). Sleep Loss and “Divergent” Thinking Ability. *Sleep*, 11, 528-536.
- Hurd, R. (2007). Fatigue. Medline plus Medical Encyclopedia. U.S. National Library of Medicine and National Institutes of Health. <http://www.nlm.nih.gov/medlineplus/ency/article/003088.htm>. Accessed 1/25/2015
- Jann, K., Dierks, T., Boesch, C., Kottlow, M., Strik, W., & Koenig, T. (2009). BOLD Correlates of EEG Alpha Phase-Locking and the fMRI Default Mode Network. *Neuroimage*, 45, 903-916. <http://dx.doi.org/10.1016/j.neuroimage.2009.01.001>
- Jann, K., Kottlow, M., Dierks, T., Boesch, C., & Koenig, T. (2010). Topographic Electrophysiological Signatures of fMRI Resting State Networks. *PLoS ONE*, 5, e12945. <http://dx.doi.org/10.1371/journal.pone.0012945>
- Kennedy, J. L. (1953). Some Practical Problems of the Alertness Indicator. In W. F. Floyd, & A. T. Welford (Eds.), *Symposium on Fatigue* (pp. 149-153). Oxford: H. K. Lewis & Co.
- Klimesch, W. (1999). EEG Alpha and Theta Oscillations Reflect Cognitive and Memory Performance: A Review and Analysis. *Brain Research Reviews*, 29, 169-195. [http://dx.doi.org/10.1016/S0165-0173\(98\)00056-3](http://dx.doi.org/10.1016/S0165-0173(98)00056-3)
- Klimesch, W., Doppelmayr, M., Schimke, H., & Ripper, B. (1997). Theta Synchronization and Alpha Desynchronization in a Memory Task. *Psychophysiology*, 34, 169-176. <http://dx.doi.org/10.1111/j.1469-8986.1997.tb02128.x>
- Krishnan, V. K., Dasari, D., & Ding, L. (2014). *EEG Correlates of Fluctuation in Cognitive Performance in an Air Traffic Control Task. Technical Report DOT/FAA/AM-14/12*. Washington DC: Office of Aerospace Medicine, Federal Aviation Administration.
- Laufs, H., Krakow, K., Sterzer, P., Eger, E., Beyerle, A., Salek-Haddadi, A., & Kleinschmidt, A. (2003). Electroencephalographic Signatures of Attentional and Cognitive Default Modes in Spontaneous Brain Activity Fluctuations at Rest. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 11053-11058. <http://dx.doi.org/10.1073/pnas.1831638100>
- Lim, J., Wu, W. C., Wang, J., Detre, J. A., Dinges, D. F., & Rao, H. (2010). Imaging Brain Fatigue from Sustained Mental Workload: An ASL Perfusion Study of the Time-on-Task Effect. *Neuroimage*, 49, 3426-3435. <http://dx.doi.org/10.1016/j.neuroimage.2009.11.020>
- Lorist, M. M., Boksem, M. A., & Ridderinkhof, K. R. (2005). Impaired Cognitive Control and Reduced Cingulate Activity During Mental Fatigue. *Cognitive Brain Research*, 24, 199-205. <http://dx.doi.org/10.1016/j.cogbrainres.2005.01.018>
- Makeig, S., & Inlow, M. (1993). Lapses in Alertness: Coherence of Fluctuations in Performance and EEG Spectrum. *Electroencephalography and Clinical Neurophysiology*, 86, 23-35. [http://dx.doi.org/10.1016/0013-4694\(93\)90064-3](http://dx.doi.org/10.1016/0013-4694(93)90064-3)
- Mast, T. M., & Heimstra, N. W. (1964). Effects of Fatigue on Vigilance Performance. *Journal of Engineering Psychology*, 3, 73-79.
- Mecklinger, A., Kramer, A. F., & Strayer, D. L. (1992). Event Related Potentials and EEG Components in a Semantic Memory Search Task. *Psychophysiology*, 29, 104-119. <http://dx.doi.org/10.1111/j.1469-8986.1992.tb02021.x>
- Montgomery, L. D., Montgomery, R. W., & Guisado, R. (1995). Rheoencephalographic and Electroencephalographic Measures of Cognitive Workload: Analytical Procedures. *Biological Psychology*, 40, 143-159. [http://dx.doi.org/10.1016/0301-0511\(95\)05117-1](http://dx.doi.org/10.1016/0301-0511(95)05117-1)
- Murta, T., Leite, M., Carmichael, D. W., Figueiredo, P., & Lemieux, L. (2015). Electrophysiological Correlates of the BOLD Signal for EEG-Informed fMRI. *Human Brain Mapping*, 36, 391-414. <http://dx.doi.org/10.1002/hbm.22623>
- Neri, D. F., Shappell, S. A., & DeJohn, C. A. (1992). Simulated Sustained Flight Operations and Performance, Part 1: Effects of Fatigue. *Military Psychology*, 4, 137-155. [http://dx.doi.org/10.1207/s15327876mp0403\\_2](http://dx.doi.org/10.1207/s15327876mp0403_2)

- Nyenhuis, D. L., Stern, R. A., Yamamoto, C., Luchetta, T., & Arruda, J. E. (1997). Standardization and Validation of the Visual Analogue Mood Scales. *Clinical Neuropsychologist*, *11*, 407-415. <http://dx.doi.org/10.1080/13854049708400470>
- Pacific Development and Technology, LLC (2012). APECSgui Installation Guide and User Manual Version 1.0. Accessed 9 February 2015. <http://publications.neurodia.com/PDT-APECSgui-Software-Manual-V1.0.1.pdf>
- Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical Splines for Scalp Potential and Current Density Mapping. *Electroencephalography and Clinical Neurophysiology*, *72*, 184-187. [http://dx.doi.org/10.1016/0013-4694\(89\)90180-6](http://dx.doi.org/10.1016/0013-4694(89)90180-6)
- Pfurtscheller, G., & Lopes da Silva, F. H. (2005). Event-Related Desynchronization (ERD) and Event-Related Synchronization (ERS). In E. Niedermeyer, & F. H. Lopes da Silva (Eds.), *Electroencephalography, Basic Principles, Clinical Applications and Related Fields* (pp. 1003-1016). Philadelphia: Lippincott Williams & Wilkins.
- Rosipal, R., & Trejo, L. J. (2001). Kernel Partial Least Squares in Reproducing Kernel Hilbert Space. *Journal of Machine Learning Research*, *2*, 97-123.
- Rosipal, R., Girolami, M., Trejo, L. J., & Cichocki, A. (2001). Kernel PCA for Feature Extraction and De-Noising in Nonlinear Regression. *Neural Computing & Applications*, *10*, 231-243. <http://dx.doi.org/10.1007/s521-001-8051-z>
- Rosipal, R., Trejo, L. J., & Nunez, P. L. (2009). Application of Multi-Way EEG Decomposition for Cognitive Workload Monitoring. In V. E. Vinzi, M. Tenenhaus, & R. Guan (Eds.), *Proceedings of the 6th International Conference on Partial Least Squares and Related Methods* (pp. 145-149). Beijing.
- Rosipal, R., Trejo, L. J., & Zaidel, E. (2013). Atomic Decomposition of EEG for Mapping Cortical Activation. *Tensor Methods for Machine Learning Workshop*, Prague, 27 September 2013.
- Rosipal, R., Trejo, L. J., & Matthews, B. (2003). Kernel PLS-SVC for Linear and Nonlinear Classification. In T. Fawcett, & N. Mishra (Eds.), *Proceedings of the Twentieth International Conference on Machine Learning* (pp. 640-647). Washington DC, 21-24 August 2003.
- Russo, M. B., Stetz, M. C., & Thomas, M. L. (2005). Monitoring and Predicting Cognitive State and Performance via Physiological Correlates of Neuronal Signals. *Aviation, Space, and Environmental Medicine*, *76*, C59-C63.
- Sagberg, F., Jackson, P., Kruger, H.-P., Muzer, A., & Williams, A. (2004). Fatigue, Sleepiness and Reduced Alertness as Risk Factors in Driving. *TOI Report 739/2004*, Institute of Transport Economics, Oslo.
- Shen, A., Li, X., Ong, C., Shao, S., & Wilder-Smith, P. V. (2008). EEG-Based Mental Fatigue Measurement Using Multi-Class Support Vector Machines with Confidence Estimate. *Clinical Neurophysiology*, *119*, 1524-1533. <http://dx.doi.org/10.1016/j.clinph.2008.03.012>
- Shingledecker, C. A., & Holding, D. H. (1974). Risk and Effort Measures of Fatigue. *Journal of Motor Behavior*, *6*, 17-25. <http://dx.doi.org/10.1080/00222895.1974.10734975>
- Šípoš, I. (1971). Verbal Vigilance Performance and Visual Fatigue. *Studia Psychologica*, *13*, 141-142.
- Smit, A. S., Eling, P. A. T. M., & Coenen, A. M. L. (2004). Mental Effort Causes Vigilance Decrease due to Resource Depletion. *Acta Psychologica*, *115*, 35-42. <http://dx.doi.org/10.1016/j.actpsy.2003.11.001>
- Smith, M. E., Gevins, A., Brown, H., Karnik, A., & Du, R. (2001). Monitoring Task Loading with Multivariate EEG Measures during Complex Forms of Human-Computer Interaction. *Human Factors*, *43*, 366-380. <http://dx.doi.org/10.1518/001872001775898287>
- Stern, R. A. (1997). *Visual Analogue Mood Scales*. Odessa, FL: Psychological Assessment Resources.
- Thayer, R. E. (1986). Activation-Deactivation Adjective Check List: Current Overview and Structural Analysis. *Psychological Reports*, *58*, 607-614. <http://dx.doi.org/10.2466/pr0.1986.58.2.607>
- Thayer, R. E. (1989). *The Biopsychology of Mood and Arousal*. New York: Oxford University Press.
- Trejo, L. J., Matthews, R., & Allison, B. Z. (2007). Experimental Design and Testing of a Multimodal Cognitive Overload Classifier. In D. D. Schmorrow, D. M. Nicholson, J. M. Drexler, & L. M. Reeves (Eds.), *Foundations of Augmented Cognition* (4th ed., pp. 13-22). Arlington, VA: Strategic Analysis, Inc.
- Trejo, L. J., Rosipal, R., & Matthews, B. (2006). Brain-Computer Interfaces for 1-D and 2-D Cursor Control: Designs Using Volitional Control of the EEG Spectrum or Steady-State Visual Evoked Potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *14*, 225-229. <http://dx.doi.org/10.1109/TNSRE.2006.875578>
- Trejo, L., Knuth, K., Prado, R., Rosipal, R., Kubitz, K. Kochavi, R., Matthews, B., & Zhang, Y. (2007). EEG-Based Estimation of Mental Fatigue: Convergent Evidence for a Three-State Model. *HCI International 2007 and Augmented Cognition International Conference*, Beijing, 22-27 July 2007. In D.D. Schmorrow, & L.M. Reeves (Eds.), *Augmented Cognition, HCII 2007, LNAI 4565* (pp. 201-211). New York: Springer LNCS.
- Trejo, L. J., & Shensa, M. J. (1999). Feature Extraction of Event-Related Potentials Using Wavelets: An Application to Human Performance Monitoring. *Brain and Language*, *66*, 89-107. <http://dx.doi.org/10.1006/brln.1998.2026>
- Trejo, L. J., Rosipal, R., & Nunez, P. L. (2010). Advanced Physiological Estimation of Cognitive Status. *The 27th Army Science Conference*, Orlando, 29 November-2 December 2010.

- U.S. Congress, Office of Technology Assessment (1991). *Biological Rhythms: Implications for the Worker*, OTA-BA-463. Washington DC: U.S. Government Printing Office.
- van der Hulst, M., Meijman, T., & Rothengatter, T. (2001). Maintaining Task Set under Fatigue: A Study of Time-on-Task Effects in Simulated Driving. *Transportation Research Part F: Traffic Psychology & Behavior*, 4, 103-118. [http://dx.doi.org/10.1016/S1369-8478\(01\)00017-1](http://dx.doi.org/10.1016/S1369-8478(01)00017-1)
- van der Linden, D., Frese, M., & Meijman, T. F. (2003). Mental Fatigue and the Control of Cognitive Processes: Effects on Perseveration and Planning. *Acta Psychologica*, 113, 45-65. [http://dx.doi.org/10.1016/S0001-6918\(02\)00150-6](http://dx.doi.org/10.1016/S0001-6918(02)00150-6)
- van der Linden, D., Frese, M., & Sonnentag, S. (2003). The Impact of Mental Fatigue on Exploration in a Complex Computer Task: Rigidity and Loss of Systematic Strategies. *Human Factors*, 45, 483-494. <http://dx.doi.org/10.1518/hfes.45.3.483.27256>
- Vidulich, M. A., Stratton, M., Crabtree, M., & Wilson, G. (1994). Performance-Based and Physiological Measures of Situational Awareness. *Aviation, Space, and Environmental Medicine*, 65, A7-A12.
- Wallerius, J., Trejo, L. J., Matthews, R., Rosipal, R., & Caldwell, J. A. (2005). Robust Feature Extraction and Classification of EEG Spectra for Real-Time Classification of Cognitive State. *Proceedings of 11th International Conference on Human Computer Interaction*, Las Vegas.
- Wilson, G. F., & Fisher, F. (1991). The Use of Cardiac and Eye Blink Measures to Determine Flight Segment in F4 Crews. *Aviation, Space and Environmental Medicine*, 62, 959-962.
- Wilson, G. F., & Fisher, F. (1995). Cognitive Task Classification Based upon Topographic EEG Data. *Biological Psychology*, 40, 239-250. [http://dx.doi.org/10.1016/0301-0511\(95\)05102-3](http://dx.doi.org/10.1016/0301-0511(95)05102-3)
- Wilson, G. F., & Russell, C. A. (2003a). Operator Functional State Classification Using Multiple Psychophysiological Features in an Air Traffic Control Task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 45, 381-389. <http://dx.doi.org/10.1518/hfes.45.3.381.27252>
- Wilson, G. F., & Russell, C. A. (2003b). Real-Time Assessment of Mental Workload Using Psychophysiological Measures and Artificial Neural Networks. *Human Factors*, 45, 635-643. <http://dx.doi.org/10.1518/hfes.45.4.635.27088>
- Winningham, M. L. (1996). Fatigue. In S. L. Groenwald, M. Hansen Frogge, M. Goodman, & C. H. Yarbro (Eds.), *Cancer Symptom Management* (pp. 42-54). Sudbury, MA: Jones and Bartlett.
- Zaidel, E., Rosipal, R., Hill, A., Fernandes, N., Akbarut, R., Noh, S., & Trejo, L. J. (2013). Modeling EEG-Band Neurofeedback: Modulating Internal States without Conditioning of EEG Sources. *ISNR 21st Annual Conference*, Dallas, 18-22 September 2013.