Influence of Task Combination on EEG Spectrum Modulation for Driver Workload Estimation

Shengguang Lei and **Matthias Roetting,** Berlin Institute of Technology, Berlin, Germany

Objective: This study investigates the feasibility of using a method based on electroencephalography (EEG) for deriving a driver's mental workload index.

Background: The psychophysiological signals provide sensitive information for human functional states assessment in both laboratory and real-world settings and for building a new communication channel between driver and vehicle that allows for driver workload monitoring.

Methods: An experiment combining a lane-change task and *n*-back task was conducted. The task load levels were manipulated in two dimensions, driving task load and working memory load, with each containing three task load conditions.

Results: The frontal theta activity showed significant increases in the working memory load dimension, but differences were not found with the driving task load dimension. However, significant decreases in parietal alpha activity were found when the task load was increased in both dimensions. Task-related differences were also found. The driving task load contributed more to the changes in alpha power, whereas the working memory load contributed more to the changes in theta power. Additionally, these two task load dimensions caused significant interactive effects on both theta and alpha power.

Conclusion: These results indicate that EEG technology can provide sensitive information for driver workload detection even if the sensitivities of different EEG parameters tend to be task dependent.

Application: One potential future application of this study is to establish a general driver workload estimator that uses EEG signals.

Keywords: electroencephalography, operator functional states, driver mental states, psychophysiological measures, *n*-back, lane-change task

Address correspondence to Shengguang Lei, Technical University of Berlin, Franklinstr 28/29, FR 2-7/1, Berlin 10587, Germany; e-mail: sle@mms.tu-berlin.de.

HUMAN FACTORS
Vol. 53, No. 2, April 2011, pp. 168–179.
DOI: 10.1177/0018720811400601.

INTRODUCTION

With the spread of in-vehicle technologies (IVTs), such as navigation systems and cellular phones, the driver often engages in multiple tasks unrelated to vehicle control and navigation (Lenneman & Backs, 2009). In such situations, driver mental overload may occur, especially if confronted with complex driving conditions (e.g., high traffic density or poor weather), and the likelihood of driving error increases (De Waard, 1996). Hence, the evaluation of driver workload is important for traffic safety research.

Driver workload has been examined with various methods in recent decades, such as performance evaluation, subjective reporting, and psychophysiological measurements. Performance measurement is used to examine the detrimental effect of various side tasks and devices, such as navigation systems (e.g., Tsimhoni, Smith, & Green, 2004) and cellular phones (e.g., Collet, Guillot, & Petit, 2010; Strayer, Drews, & Crouch, 2006) on the driving task. Subjective assessment tools, such as the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988), are used to provide workload self-evaluation through responses to certain questionnaires. Specifically, the Driver Activity Load Index (DALI) was developed as a subjective method to evaluate driver workload (Pauzie, 2008). However, performance measurement demonstrates occasional insensitivity (Lenneman & Backs, 2009), and subjective evaluation does not measure time-varying qualities and is often influenced by events toward the end of immersion at the time of questionnaire administration (Insko, 2003).

Psychophysiological measures, for example, electroencephalogram (EEG), electrocardiogram (ECG), and pupil dilation, address shortcomings in performance and subjective measurements and are found to be robust candidates for operators workload evaluation (Brookhuis & De

Waard, 2010; Gevins et al., 1998; Gevins, Smith, McEvoy, & Yu, 1997; Kramer, 1991; Wilson & Russell, 2003). Unlike performance and subjective measurements, psychophysiological measures offer continuous observation in high time resolution (e.g., in milliseconds) and can be collected without intruding into the operator's task (Kramer, 1991; Wilson & Russell, 2007).

The EEG signal is a representation of the brains electrical activity recorded from electrodes placed on the scalp. It has been used to assess operators workload for many years in both laboratory (Berka et al., 2007; Gundel & Wilson, 1992; Lei, Welke, & Roetting, 2009) and applied settings (Kohlmorgen et al., 2007; Wilson, 2001, 2002). The EEG spectral components, for example, theta (4–8 Hz) and alpha (8–12 Hz), are used to determine activity levels during different cognitive activities. The majority of previous findings consistently indicate that increased workload leads to increased frontal theta (fro-theta) activity and decreased parietal alpha activity (Gevins et al., 1997, 1998; Gundel & Wilson, 1992; Smith, Gevins, Brown, Karnik, & Du, 2001; Sterman, Mann, Kaiser, & Suyenobu, 1994; Yamamoto & Matsuoka, 1990).

EEG spectrum modulation has also been introduced to investigate driver workload in various driving conditions (Brookhuis & De Waard, 1993; Hagemann, 2008; Kohlmorgen et al., 2007). Brookhuis and De Waard (1993) used an energy parameter ([theta + alpha] / beta) to measure participants activation during on-the-road driving experiments. More recently, Kohlmorgen and colleagues (2007) outlined an EEG-based system for detecting driver mental workload in real traffic conditions. They classified driver workload into high and low conditions with various EEG spectrum features. The result was used immediately to modulate the workload induced by the influx of information from the cars electronic systems. They showed that a system as such was beneficial for improving drivers overall task performance.

Besides EEG components, the ECG signal, which reflects heart activity, has been introduced for driver workload assessment (Brookhuis & De Waard, 1993, 2010; Lenneman & Backs, 2009; Mehler, Reimer, Coughlin, & Dusek, 2009). The ECG involves several parameters, such as heart rate (HR; the number of heartbeats within a fixed period of time), the interbeat interval (IBI), and heart rate variability (HRV; changes of the interval between heartbeats in either time or frequency domain; see Kramer, 1991). An increase in HR and a decrease in HRV are expected when more mental effort is required (see Brookhuis & De Waard, 2010; Mulder, De Waard, & Brookhuis, 2004). In the HRV frequency domain, low frequency (LF; 0.04–0.15 Hz), high frequency (FH; 0.15–0.4 Hz), and their ratio (LF/HF) are all sensitive to changes in the operator's workload (Kamada, Miyake, Kumashiro, Monou, & Inoue, 1992; Murai, Hayashi, Nagata, & Inokuchi, 2004; Wilson, 2002).

Modulation of the EEG spectrum in simple tasks whereby workload levels were manipulated in a single dimension (e.g., working memory load), has been systematically investigated (Gevins et al., 1997, 1998; Gundel & Wilson, 1992). There are also other studies involving relatively complex tasks (Smith et al., 2001; Wilson & Russell, 2003). However, most of these either investigate EEG spectrum modulation in a general manner or directly use machine learning methods to classify workload levels with multiple EEG variables. There is little evidence of a clear comparison between the similarities and differences in the EEG spectrum modulation induced by different workload dimensions. Questions such as how these workload subresources contribute to the general workload and what effect will emerge after these subresources are combined should be addressed.

An objective of this study is to determine whether previous findings are reproducible in driving contexts. Another objective is to investigate the effects of different workload dimensions and their combinations on EEG spectrum variation. Unlike Kohlmorgen et al.'s (2007) study, we focus on the analysis of changes in the EEG spectrum induced by the task load rather than use a machine learning method classifying the workload. A simulated driving task, the lane-change task (LCT; Burns, Trbovich, McCurdie, & Harbluk, 2005; Mattes, 2003), combined with a secondary working memory task, the *n*-back task (Kirchner, 1958), was adopted in this experiment. Besides EEG, other

Figure 1. Experiment set-up: A driving simulator.

variables, including NASA-TLX, ECG, *n*-back performance, and reaction time, were recorded to determine the feasibility of using EEG to index driver workload.

METHOD

Participants

Overall, 26 participants (19 males, 7 females) participated in this study. They all ranged in age from 21 to 33 with a mean of 27.8 (standard deviation = 2.96 years). Of the participants, 15 had a valid driving license. All individuals were reported to be free of illness and medication. None had prior experience with either the LCT or the *n*-back task. Data from 2 participants were excluded because of incomplete recordings during the experiment attributable to simulator sickness. All participants received cash remuneration for their participation.

Experiment Apparatus

The experiment was conducted with the use of a driving box (Figure 1). Using a projector, we projected the driving scene onto the wall approximately 1 m in front of the driving box. The *n*-back digits were visually presented to the participants with projections overlaid on the driving scene by another projector (Figure 2). The driving box had an adjustable driving seat and buttons on the steering wheel that allowed participants to react to the *n*-back task.

Brain activity was recorded with 32 Ag/AgCl impedance-optimized electrodes (ActiCap, Brain Products, Germany), referenced to the nasion, sampled at 1000 Hz and wide band filtered (0.5–70 Hz), and placed according to the international 10-20 system. One channel of the ECG was used to collect the heartbeat information via two bipolar electrodes. One electrode was attached at the upper breastbone and the

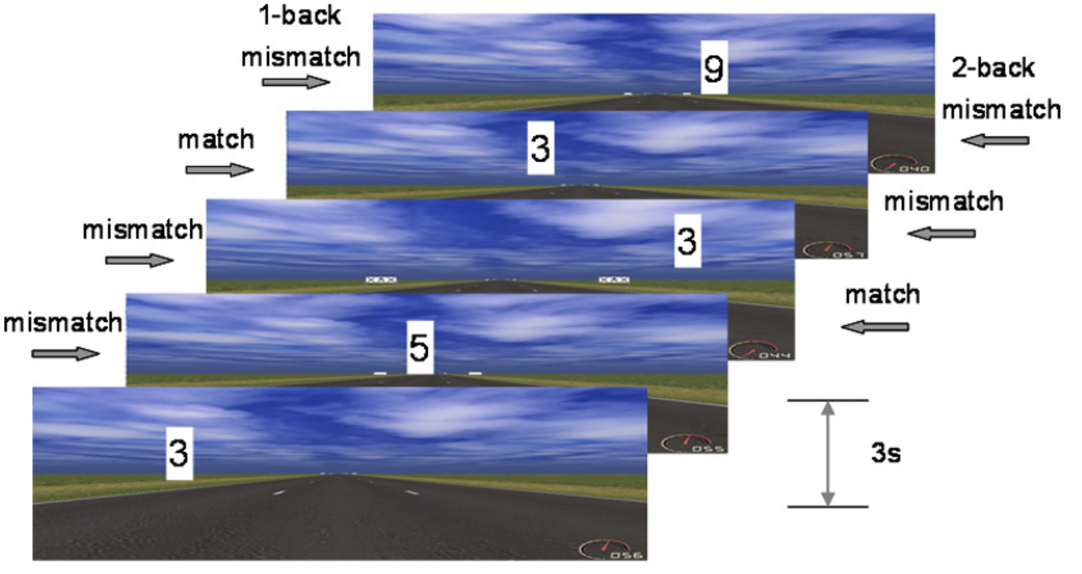

Figure 2. Overlaid projections of the *n*-back tasks and driving scene. The digits of the *n*-back tasks were randomly allocated to the driving scene.

other around the lowest rib on the left side of the body. ECG was sampled at 1000 Hz and recorded with EEG as one extended channel with the use of the software Brain Vision Recorder from Brain Products.

The signals from the LCT, *n*-back, and EEG were synchronized with the use of a self-developed tool based on Labview (National Instruments, USA). This tool automatically triggered the *n*-back program when detecting the start of each LCT track and shut it down at the end of each track.

Tasks

LCT. The LCT was initiated by the project ADAM (Advanced Driver Attention Metrics) as a standardized methodology for evaluating attentional demands associated with performing in-vehicle tasks during driving. In the LCT, the driving road consists of three lanes, and participants are asked to repeatedly perform lane changes when prompted by road signs without other vehicles or pedestrians present. The LCT consists of 10 tracks; each track is approximately 3 km (approximately 3 min if driving at 60 km/h) and includes 18 randomly sequenced road signs (indicating to which lane the participant should change) plus a start sign. In the present study, the participant was asked to fully step on the gas pedal to travel at a constant maximum speed configured in advance.

The quality of these lane changes can be evaluated by the difference (based on mean deviation) between a normative lane change path and the drivers actual lane change path, influenced by the drivers ability to detect and respond to the road signs as well as to maintain lateral control. In this study, we used a linear global normative path for all participants in all driving conditions.

N-*back task.* The *n*-back task was first introduced by Kirchner in 1958 and is commonly used in neuroimaging to stimulate brain activity to test working memory capacity. The participant is presented with a sequence of stimuli (in this study, the stimuli were digits), and the task consists of indicating whether the current stimulus matches the one from *n* steps used earlier in the sequence. The load factor *n* can be adjusted to make the task more or less difficult. In this study, the digital stimuli were presented every 3 s for a duration of 1 s. Each time, the digit appeared in a random location on the driving scene. Only 1-back and 2-back were used. Participants reacted to *n*-back tasks by pressing the left or right button on the steering wheel to indicate match or mismatch of the digits, respectively.

Task load manipulation. A 3 × 3 two-factor within-subject design was used in this study. Task load levels were manipulated with two dimensions, driving task load and working memory load. The driving task load dimension comprised three conditions: no driving ("base"; participants passively watched the driving performed by the experimenter at a speed of 75 km/h), slow driving ("slow"; driving speed was set at 75 km/h while participants used the steering wheel to perform lane changes), and fast driving ("fast"; driving speed was set at 100 km/h, and participants used the steering wheel to perform lane changes). The working memory load dimension also contained three conditions: no *n*-back (N0; digits were still presented to participants but required no reaction from them), 1-back (N1), and 2-back (N2). The combination of driving conditions and *n*-back modes resulted in a total of nine task load levels.

Procedures

Participants first filled out a form with their personal information (age, driving experience, illness and medication situation, etc.) before reading an experiment introduction. After the electrode preparation (approximately 20 min), they had a 20-min practice session with all nine task conditions. Then, they were required to evaluate the NASA-TLX weights of six workload contributors. Afterward, they were asked to perform three randomly organized sessions (base, slow, and fast), with a 5-min break after each session. Both base and slow sessions comprised 9 randomized driving trials, 3 trials on each *n*-back condition, lasting 2.5 min per trial. The fast trials lasted approximately 2 min each. To keep the same duration for each experimental condition, the fast session comprised 12 trials (randomized as well), 4 trials per *n*-back condition. Generally, each session lasted 25 min. During each session break, participants were asked to report NASA-TLX ratings. The whole experiment lasted approximately 2 hr and 40 min.

Data Analysis

In total, eight parameters, that is, fro-theta power, parietal alpha power (par-alpha), subjective load, HR, HRV, LCT performance, *n*-back performance, and reaction time, were extracted to examine variations with the task load.

EEG analysis. We performed an EEG analysis using EEGLAB 6.03, a freely available open-source toolbox that runs on Matlab 7.3.0 (see Delorme & Makeig, 2004). We digitally filtered EEG data using a band pass filter (pass band 1–40 Hz) to minimize drifts and line noises. Because EEG data involve plenty of eye movement artifacts, independent components analysis (Delorme & Makeig, 2004; Makeig, Bell, Jung, & Sejnowski, 1996) was used for ocular artifacts removal. The EEG data were then segmented into 10-s epochs with 50% overlay (short-term variation of EEG parameters is also our concern although not included in this article). An FFT analysis was then used to estimate the spectrum for each epoch, and the relative power density spectrum (percentage of total power of frequency range 4–30 Hz) was calculated for each epoch. Frequency bands, theta and alpha, were extracted by accumulating the power of frequency bands 4 to 8 Hz and 8 to 12 Hz. This way, an average of 100 data points in theta and alpha power were extracted for each task condition and participant. The present study concentrated only on the modulation of the fro-theta and par-alpha, extracted by averaging the theta power from five frontal electrodes (Fz, FC1, FCz, FC2, Cz) and alpha power from five parietal electrodes (PCz, P3, Pz, P4, POz). To reduce the individual variation, the z scores of theta and alpha power were calculated for 900 data points for each participant and averaged across the task conditions.

ECG analysis. A module supplied with EEGLAB software used a QRS complex detection algorithm to determine individual heartbeats from the ECG signal. HR was calculated as the number of QRS complex divided by the duration for each driving track. The IBI of successive heartbeats was then used to calculate the HRV statistics. The spectrum of HRV was estimated with a method developed by Malik et al. (1996). The ratio of energy around the LF (0.04–0.15 Hz) activity to the energy around the HF (0.15–0.4 Hz) activity was analyzed to offer an alternative variable for the workload evaluation.

Statistical analysis. We performed statistical analysis with PASW Statistics 18 (SPSS, USA).

TABLE 1: Comparison of the Means (standard deviations) of Different Variables From the Experiment for All Nine Task Conditions (Values Calculated From 24 Participants)

	Base			Slow			Fast		
Variables	N0	N1	N2	N0	N1	N2	N0	N1	N2
Subjective load (0–100)	14.8 (14.3)	25.2 (15.0)	40.0 (19.1)	19.3 (15.7)	39,9 (16.5)	56.7 (16.0)	28.2 (20.4)	45.3 (16.6)	68.5 (15.3)
N-back RT (s)	–	.78 (.11)	.97 (.17)	–	.95 (.21)	1.10 (.24)	–	0.98 (.20)	1.09 (.20)
N-back error rate (0–1)	–	0.05 (.04)	0.10 (.09)	–	.09 (.08)	.19 (.14)	–	.13 (.11)	.22 (.14)
LCT mean deviation(m)	–	–	–	1.79 (.29)	1.87 (.28)	1.89 (.29)	2.29 (.41)	2.52 (.42)	2.52 (.45)
Heart Rate	72.0 (7.4)	72.4 (7.7)	73.5 (7.6)	73.5 (8.4)	75.8 (8.5)	76.6 (8.0)	74.0 (8.9)	75.7 (8.7)	77.1 (8.3)
Heart Rate Variability (LF/HF ratio)	2.06 (1.45)	1.76 (1.63)	1.50 (1.10)	1.98 (1.85)	1.92 (1.49)	1.49 (1.07)	1.58 (1.00)	1.49 (1.04)	1.31 (1.04)

A two-way ANOVA was used to test the significance of the differences in these parameters, and then multiple comparisons were made with post hoc analysis (Bonferroni). An alpha of .05 determined the statistical significance.

RESULTS

Subjective Load

As shown in Table 1, the subjective workload showed significant increases in both dimensions: for driving task load, $F(2, 24) = 34.7$, $p < .001$; for working memory load, $F(2, 24) = 95.2$, $p < .001$. The interaction of these two dimensions was also significant, $F(4, 24) = 7.7$, $p < .001$). A post hoc test indicated that within each *n*-back level, significant differences between each pair of driving task conditions were found but with no differences between slow and fast in N1 and N2 conditions (in N0, base and slow, $p < .05$; base and fast, $p < .001$; slow and fast, $p < .01$; in N1, base and slow, $p < .001$; base and fast, $p < .001$; slow and fast, $p = .14$; in N2, base and slow, $p < .001$; base and fast, $p < .001$; slow and fast, $p = .09$). Within each driving task level, significant differences between each pair of *n*-back conditions were found (for all pairs, $p < .001$).

Task Performance

Task performances are shown in Table 1. The mean deviation in LCT showed significant increases with augmented driving task load, $F(1, 24) = 89.4$, $p < .001$, and working memory load, $F(2, 24) = 17.1$, $p < .001$. There was also a significant interaction effect, $F(2, 24) = 5.9$, $p < .01$. A post hoc test showed significant differences in mean deviation between slow and fast within each *n*-back level (in N0, slow and fast, $p < .001$; in N1, slow and fast, $p < .001$; in N2, slow and fast, $p < .001$). Within each driving task level, there were significant differences in mean deviation between N0 and N1 and between N0 and N2 but none between N1 and N2 (in slow, N0 and N1, $p < .05$; N0 and N2, $p < .05$; N1 and N2, $p = .64$; in fast, N0 and N1, $p < .001$; N0 and N2, $p < .001$; N1 and N2, $p = .88$).

N-back error rate showed significant increases with augmented driving task load, $F(2, 24) = 12.6$, $p < .001$, and working memory load, $F(2, 24) = 26.2$, $p < .001$. There was also a significant interaction effect, $F(2, 24) = 4.7$, $p < .05$. A post hoc test indicated that within each *n*-back level, significant differences were found between each pair of driving task load conditions except for pairing between slow and fast in the N2 condition (in N1, base and slow, $p < .05$; base and fast, $p < .001$; slow and fast, $p < .01$; in N2, base and slow, $p < .01$; base and fast, $p < .001$; slow and fast, $p = .06$). Significant differences between N1 and N2 conditions were found within each driving task level (in base, N1 and N2, $p < .01$; in slow, N1 and N2, $p < .001$; in fast, N1 and N2, $p < .001$).

Additionally, a significant response time delay for n-back task was found with increased driving task load, $F(2, 24) = 16.1$, $p < .001$, and working memory load, $F(1, 24) = 69.7$, $p < .001$. A significant interaction effect was also observed, $F(4, 24) = 5.8$, $p < .01$. A post hoc test indicated that for each n-back level, there were significant response time differences between base and slow and between base and fast but none between slow and fast (in N1, base and slow, $p < .001$; base and fast, $p < .001$; slow and fast, $p = .11$; in N2, base and slow, $p < .01$; base and fast, $p < .05$; slow and fast, $p = .77$). Significant differences between N1 and N2 conditions were found within each driving task level (in base, N1 and N2, $p < .001$; in slow, N1 and N2, $p < .001$; in fast, N1 and N2, $p < .001$).

HR and HRV

As shown in Table 1, a significant rise in HR was found as the task load level increased in both dimensions: for driving task load, $F(2, 24) = 12.2$, $p < .001$; for working memory load, $F(2, 24) = 31.9$, $p < .001$, with significant interaction effect, $F(4, 24) = 8.1$, $p < .001$. A post hoc test indicated that within each n-back level, there were significant differences in HR between base and slow and between base and fast but none between slow and fast (in N0, base and slow, $p < .05$; base and fast, $p < .05$; slow and fast, $p = .43$; in N1, base and slow, $p < .001$; base and fast, $p < .01$; slow and fast, $p = .87$; in N2, base and slow, $p < .001$; base and fast, $p < .001$; slow and fast, $p = .36$. Within each driving task level, there were significant differences between each pair of n-back conditions except N0 and N1 in base (in base, N0 and N1, $p = .31$; N0 and N2, $p < .01$; N1 and N2, $p < .01$; in slow, N0 and N1, $p < .001$; N0 and N2, $p < .001$; N1 and N2, $p < .05$; in fast, N0 and N1, $p < .001$; N0 and N2, $p < .001$; N1 and N2, $p < .001$).

The LF/HF ratio in HRV showed a significant negative correlation with the workload in both dimensions: for driving task load, $F(2, 24) = 3.9$, $p < .05$; for working memory load, $F(2, 24) = 8.1$, $p < .01$. There was no significant interaction between driving task and n-back task on LF/HF, $F(4, 24) = .5$, $p = .71$. A post hoc test

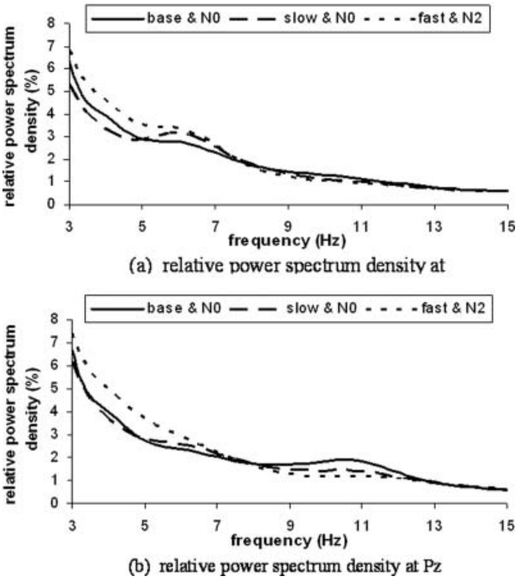

Figure 3. Relative power spectrum density (the percentages of the total power of 4–30 Hz) (a) at frontal recording site (Fz) and (b) at the parietal recording site (Pz) averaged across participants ($N = 24$).

indicated that within each n-back level, significant differences were found only between base and fast in the N0 condition ($p < .05$) and between slow and fast in the N1 condition ($p < .05$). Within each driving task level, significant differences were found only between N0 and N3 in the base condition ($p < .01$) and between N1 and N2 in the slow condition ($p < .05$).

EEG Spectrum Modulation

Figures 3a and 3b demonstrate the relative power spectrum density of three task conditions at electrode sites Fz and Pz, respectively. To illustrate a general pattern of EEG spectrum changes with task load, only three conditions (base and N0, slow and N0, and fast and N2, representing low, moderate, and high task load levels, respectively) were included in the picture. Generally, there was tendency in the power of the fro-theta frequency range to increase whenever the task load was increased (Figure 3a), whereas the parietal alpha activity strongly attenuated as the task load was increased (Figure 3b).

The results for fro-theta are shown in Figure 4. The variation of driving task load elicited no

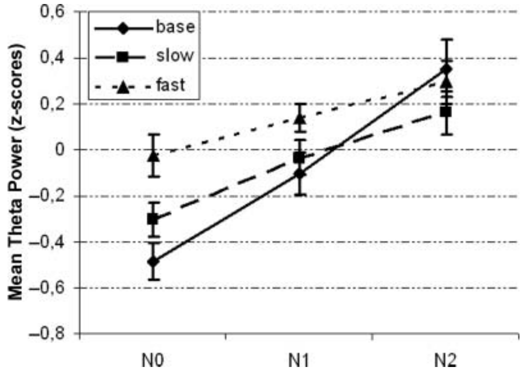

Figure 4. Mean z scores of frontal theta (4–8 Hz) for each of nine task conditions. Z scores were converted for each participant and averaged across 24 participants. The error bars show standard errors of the z scores.

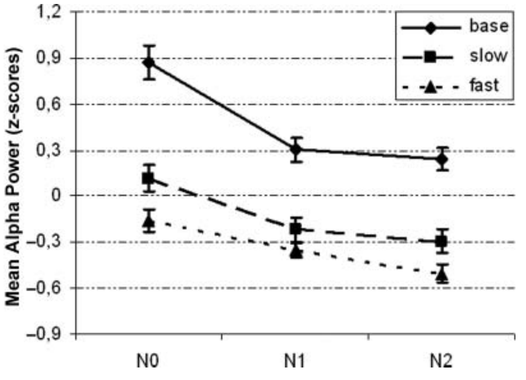

Figure 5. Mean z scores of parietal alpha (8–12 Hz) for each of nine task conditions. Z scores were converted for each participant and averaged across 24 participants. The error bars show standard errors of the z scores.

significant differences in the fro-theta, $F(2, 24) = 2.4$, $p = .10$. However, the changes in working memory load produced significant increases in the fro-theta, $F(2, 24) = 23.6$, $p < .001$, with significant interaction between them, $F(4, 24) = 4.2$, $p < .01$. A post hoc test indicated that for each driving task level, significant differences between each pair of n-back conditions were found in the base and slow condition but not in the fast condition (in base, N0 and N1, $p < .001$; N0 and N2, $p < .001$; N1 and N2, $p = .45$; in slow, N0 and N1, $p < .01$; N0 and N2, $p < .001$; N1 and N2, $p < .05$; in fast, N0 & N1, $p = .08$; N0 and N2, $p = .06$; N1 and N2, $p = .15$).

As shown in Figure 5, there were significant decreases in z scores of relative par-alpha whenever workload increased with driving task load, $F(2, 24) = 45.9$, $p < .001$, and working memory load, $F(2, 24) = 18.5$, $p < .001$, with a significant interaction effect, $F(4, 24) = 3.3$, $p < .05$). A post hoc test indicated that within each n-back level, there were significant differences between each pair of driving task load conditions except between slow and fast in the N1 condition (in N0, base and slow, $p < .001$; base and fast, $p < .001$; slow and fast, $p < .01$; in N1, base and slow, $p < .001$; base and fast, $p < .001$; slow and fast, $p = .14$; in N2, base and slow, $p <.001$; base and fast, $p < .001$; slow and fast, $p < .05$). Within each driving task level, there were significant differences between each pair of n-back conditions except between N1 and N2 in base and slow conditions (in base, N0 and N1, $p < .001$; N0 and N2, $p < .001$; N1 and N2, $p = .45$; in slow, N0 and N1, $p < .01$; N0 and N2, $p < .01$; N1 and N2, $p = .32$; in fast, N0 and N1, $p < .05$; N0 and N2, $p < .01$; N1 and N2, $p < .05$).

Correlation of EEG Parameters to Other Variables

The correlations of EEG parameters to other parameters were also investigated. As shown in Table 2, fro-theta was significantly correlated to the subjective load and HRV. However, par-alpha was significantly correlated to the subjective load, n-back error rate, HR, and HRV. Interestingly, par-alpha showed a higher correlation to HR than HRV, whereas for-theta showed a higher correlation to HRV than HR.

DISCUSSION

This study investigates EEG spectrum modulation for driver workload representation under various factors. Both objective and subjective measurements indicated that task manipulations increased the workload required for task performance. However, these measurements showed different sensitivities for workload representation. In the following section, some of these results are discussed.

TABLE 2: Correlation (Pearson *r*) of EEG Parameters to Other Variables

Variables	Subjective Load	*N*-Back Error Rate	*N*-Back RT	LCT Mean Deviation	HR	HRV (LF/HF)
fro-theta	$r = .83$	$r = .63$	$r = .70$	$r = .66$	$r = .64$	$r = -.91$
	$p < .01$	$p = .18$	$p = .12$	$p = .15$	$p = .06$	$p < .001$
par-alpha	$r = -.80$	$r = -.81$	$r = -.76$	$r = -.69$	$r = -.91$	$r = .69$
	$p < .01$	$p < .05$	$p = .07$	$p = .12$	$p < .001$	$p < .05$

EEG Spectrum Modulation With Workload

The EEG signal yielded sensitive indices for neural resources utilization and systematic variation with task loads. As presented in the Results section, the increase in working memory load resulted in an increase in fro-theta and a decrease in par-alpha, which is fairly consistent with previous findings on working memory load (Gevins et al., 1997, 1998; Gundel & Wilson, 1992; Sterman et al., 1994). Gevins and colleagues (1997, 1998) found that EEG signal in the theta range, the largest over middle-line frontal regions of the scalp, was enhanced in tasks with increased working memory load. Conversely, activities in the alpha range were attenuated by attention-demanding tasks.

Increasing driving task load resulted in a significant decrease for par-alpha but elicited no differences in the fro-theta. These findings were consistent with other driving task studies (Hagemann, 2008) and in-flight tasks (Wilson, 2001, 2002). Hagemann (2008) systematically investigated changes in alpha power with task load in driving contexts. A significant attenuation of alpha was found when participants performed simultaneously an LCT and a secondary word or tone detection task, compared with conditions of only word or tone detection. Wilson (2002) investigated the changes of EEG activity in various flight conditions. Compared with the preflight baseline, alpha band power was significantly decreased over the parietal scalp during flight tasks. However, only few experimental segments showed increased theta band activity at few scattered electrode sites with no consistent pattern evident.

Previous research suggests that fro-theta is generated in the anterior midline region of the scalp (Gevins et al., 1997; Inouye et al., 1994), which is thought to be part of an anterior brain network critical to attention control mechanisms (Posner & Peterson, 1990; Posner & Rothbart, 1992; Smith et al., 2001). The present results are consistent with the view that enhancing working memory load places high demands on frontal brain circuits in relation to attention control.

Alpha activity is considered inversely proportional to the fraction of cortical neurons recruited in a transient functional network for the purpose of task performance (Smith et al., 2001). The modulation in the magnitude of alpha power is probably driven through the oscillating synchrony of neuron groups, in which individual cells act either as resonation or oscillation (Lopes da Silva, Vos, Mooibreck, & Van Rotterdam, 1980; Smith et al., 2001). When the brain is at relative rest, a high proportion of alpha generators comes to oscillate in phases, yielding a large alpha rhythm. As task demands increase, different regions of the cortex may be recruited in the transient function network, with decline in the overall proportion of local alpha generators that passively oscillate in synchrony with the reduction in alpha power (Smith et al., 2001). The observation of reduced alpha power, with both increased working memory load and driving task load, is consistent with this alpha generation hypothesis.

Specifically, statistical results show that working memory load elicited more significant differences in theta power than those elicited by driving task load. However, driving task load produced more differences in alpha power than those produced by working memory load. These results reveal a task-dependent workload effect on the modulation of EEG activity in either the sensitivity or activated location of the scalp, attributable to the recruitment of neural resources linked with individual regions of the brain. Metabolic studies suggest that working memory involves a functional network linking regions of the prefrontal cortex with posterior association cortices (Frisk & Milner, 1990; Gevins & Smith, 2000). On the other hand, brain imaging studies on neural correlation of driving revealed the activation of cerebellar and occipital areas related to visuomotor integration (Calhoun et al., 2002; Walter et al., 2001).

Interestingly, the combination of the two tasks leads to a significant interaction effect on both fro-theta and par-alpha activity. As shown in Figures 4 and 5, the degree of changes in fro-theta and par-alpha decreased with the combined task load. A reasonable assumption for this phenomenon is that both theta and alpha power have a nonlinear relation with a decreased slope attributable to workload increase. However, as far as we know, there is no hard evidence to support this hypothesis.

Other Variables and Their Correlation to EEG Parameters

Subjective ratings augmented with both workload dimensions while showing significant differences in each pairwise comparison. This result indicates that the manipulation of task loads was successful. The detriments of performance are consistent with Wickens's (2002) multiple resource theory. Wickens suggested that attention resources were limited and that resource structure can be described by four different dichotomies: two states of processing (perceptual-central and response), two modalities of perception (auditory and visual), two codes of processing (spatial and verbal), and two aspects of visual processing (focal and ambient). Tasks involving shared resources lead to a decline in task performance when resources are not adequate to meet task demands. In this study, both tasks involve common visual perception modality. The dual tasks elicited competence in visual resources, which produces impairment in task performance.

The results of HR data reinforce previous work indicating that HR can be sensitive to workload in driving environments (Lenneman & Backs, 2009; Mehler et al., 2009). That HRV LF/HF decreased with task load is also consistent with prior studies (Kamada et al., 1992; Murai et al., 2004; Wilson, 2002). The simple HR measure seems to demonstrate a more robust sensitivity than the complex HRV index. And this presumption is also supported by other studies in which HRV was not as sensitive to the varied cognitive demands of flight as other psychophysiological variables, such as HR and electrodermal activity (EDA) (Veltman & Gaillard, 1996; Wilson, 2002).

It is also interesting to compare the robustness of ECG and EEG parameters. Generally, the sensitivity of EEG parameters to the index workload was higher than HR and HRV. Both fro-theta and par-alpha demonstrated significant differences for each pairwise comparison with working memory load and driving task load, respectively. However, neither HR nor HRV could completely distinguish three driving task conditions in pairwise comparisons. This robustness of EEG observation is also supported by Brookings, Wilson, and Swain (1996). However, the sensitivity of these psychophysiological parameters seems task dependent, whereas each of the physiological parameters provides unique information concerning cognitive load (Wilson, 2002). For example, in this study, HR could distinguish all three *n*-back conditions, whereas fro-theta was not as sensitive in distinguishing driving conditions. Doyle et al. (2009) also suggested that EEG measures exhibit less sensitivity than HR when distinguishing cognitive load during a satellite management decision-training task.

The finding that both theta and alpha are highly correlated to subjective load emphasizes the fact that EEG parameters can be used to represent workload. What is interesting and surprising is that fro-theta demonstrates a higher correlation to HRV, whereas alpha power demonstrates a higher correlation to HR. Unfortunately, as far as we know, the correlations of EEG to ECG parameters have not been systematically investigated.

CONCLUSION AND APPLICATION

This study investigated driver workload in two dimensions, driving task load and working memory load. The results indicate that enhanced working memory load induced an increase in theta power and a decrease in alpha power, and an increased driving task load led to a decrease in alpha power. Additionally, when the two task loads were combined, there were significant interactive effects on the changes to both theta and alpha power. However, task-related differences were also observed. For instance, variation in working memory load contributed more to changes in fro-theta power and less to changes in par-alpha power, compared with driving task load. These results indicate that EEG technology can provide sensitive information for driver workload detection. However, the sensitivities of different EEG parameters tend to be task dependent. A potential application of this study is in providing some theoretical foundation for the establishment of a general metric for driver workload estimation.

ACKNOWLEDGMENTS

The authors would like to thank Christian Kothe for the *n*-back software implementation, Mario Lasch and Stefan Damke for the technical support during the experiment setup, and Marco Pedrotti and John Romanishin for reviewing this paper.

KEY POINTS

- Previous findings, of increased frontal theta power and decreased parietal alpha power with increased workload, were reproduced in the driving task.
- Task-related differences were found in this study, namely, that working memory load contributed more to increasing theta power and less toward reducing alpha power, compared with the driving task load.
- The combination of these two types of task loads had a significant interaction effect on the changes to both theta and alpha power. A reduction in changes to EEG parameters was also observed when task load was increased.

REFERENCES

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine*, *78*(5, Suppl.), B231–B244.

Brookhuis, K. A., & De Waard, D. (1993). The use of psychophysiology to assess driver status. *Ergonomics*, 36, 1099–1110.

Brookhuis, K. A., & De Waard, D. (2010). Monitoring drivers mental workload in driving simulators using physiological measures. *Accident Analysis & Prevention*, *42*, 893–903.

Brookings, J. B., Wilson G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology*, *42*, 361–377.

Burns, P. C., Trbovich, P. L., McCurdie, T., & Harbluk, J. L. (2005). Measuring distraction: Task duration and the lane-change test (LCT). In *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1980–1983). Santa Monica, CA: Human Factors and Ergonomics Society.

Calhoun, V. D., Pekar, J. J., McGinty, V. B., Adali, T., Watson, T. D., & Godfrey, G. D. (2002). Different activation dynamics in multiple neural systems during simulated driving. *Human Brain Mapping*, *16*, 158–167.

Collet, C., Guillot, A., & Petit, C. (2010). Phoning while driving I: A review of epidemiological, psychological, behavioural and physiological studies. *Ergonomics*, 53, 589–601.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *134*, 9-21.

De Waard, D. (1996). *The measurement of drivers mental workload* (PhD thesis). University of Groningen, Traffic Research Centre, Haren, Netherlands.

Doyle, M. J., Gould, J. E., Kass, S. J., Raj, A. K., Andrasik, F., & Higgins, J. P. (2009). Psychophysiological correlates of cognitive workload during a satellite management decision-training task. In *Proceedings of Human Systems Integration Symposium.* Annapolis, MD. Retrieved from http://navalengineers.net/Proceedings/HSIS2009/Papers/Doyle_Gould_Kass_Raj_Andrasik-Higgins.pdf.

Frisk, V., & Milner B. (1990). The relationship of working memory to the immediate recall of stories following unilateral temporal or frontal lobectomy. *Neuropsychologia*, *28*, 121–135.

Gevins, A., & Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex*, *10*(9), 829–839.

Gevins, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S., & Du, R. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors*, *40*, 79–91.

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.

Gundel, A., & Wilson, G. F. (1992). Topographical changes in the ongoing EEG related to the difficulty of mental tasks. *Brain Topography*, *5*, 17–25.

Hagemann, K. (2008). *The alpha band as an electrophysiological indicator for internalized attention and high mental workload in real traffic driving* (Dissertation). Universitaet Duesseldorf, Germany.

Hart, S. G., & Staveland, L. E. (1988). Development of the NASA TLX: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam, Netherlands: North Holland Press.

Inouye, T., Shinosaki, K., Iyama, A., Matsumoto, Y., Toi, S., & Ishihara, T. (1994). Potential flow of frontal midline theta activity during a mental task in the human electroencephalogram. *Neuroscience Letters*, *169*, 145–148.

Insko, B. E. (2003). Measuring presence: Subjective, behavioral and physiological methods. In G. Riva, F. Davide, & W. A. IJsselsteijn (Eds.), *Being there: Concepts, effects and measurement of user presence in synthetic environments* (pp. 109–119). Amsterdam, Netherlands: Ios Press.

Kamada, T., Miyake, S., Kumashiro, M., Monou, H., & Inoue, K. (1992). Power spectral analysis of heart rate variability in Type As and Type Bs during mental workload. *Psychosomatic Medicine*, *54*, 462–470.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*, 352-358.

Kohlmorgen, J., Dornhege, G., Braun, M. L., Blankertz, B., Müller, K. R., Curio, G., Hagemann, K., Bruns, A., Schrauf, M., & Kincses, W. E. (2007). Improving human performance in a real operating environment through real-time mental workload detection. In G. Dornhege, J. del R. Millán, T. Hinterberger, D. McFarland, & K.-R. Müller (Eds.), *Towards brain-computer interfacing* (pp. 409–422). Cambridge, MA: MIT Press.

Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. In D. Damos (Ed.), *Multiple task performance* (pp. 279–238). London, UK: Taylor & Francis.

Lei, S., Welke, S., & Roetting, M. (2009). Representation of drivers' mental workload in EEG data. In D. de Waard, J. Godthelp, F. L. Kooi, & K. A. Brookhuis (Eds.), *Human Factors*, *security and safety* (pp. 285–294). Maastricht, Netherlands: Shaker.

Lenneman, J. K., & Backs, R. W. (2009). Cardiac autonomic control during simulated driving with a concurrent verbal working memory task. *Human Factors*, *53*, 404–418.

Lopes da Silva, F. H., Vos, J. E., Mooibreck, J., & Van Rotterdam, A. (1980). Relative contributions of intracortical and thalamocortical processes in the generation of alpha rhythms, revealed by partial coherence analysis. *Electroencephalography and Clinical Neurophysiology*, *50*, 449–456.

Makeig, S., Bell, A. J., Jung, T. P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. *Advances in Neural Information Processing Systems*, *8*, 145–151.

Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., & Schwartz, P. J. (1996). Heart rate variability: Standards of measurement, physiological interpretation and clinical use. *Circulation*, *93*, 1043–1065.

Mattes, S. (2003, October). *The lane change task as a tool for driver distraction evaluation*. Paper presented at the IHRA-ITS Workshop on Driving Simulator Scenarios, Dearborn, MI.

Mehler, B., Reimer, B., Coughlin, J. F., & Dusek, J. A. (2009). The impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. *Transportation Research Record*, *2138*, 6–12.

Mulder, L. J. M., De Waard, D., & Brookhuis, K. A. (2004). Estimating mental effort using heart rate and heart rate variability. In N. Stanton, A. Hedge, H. W. Hendrick, K. A. Brookhuis, & E. Salas (Eds.), *Handbook of ergonomics and human factors methods* (pp. 201–208). London, UK: Taylor & Francis.

Murai, K., Hayashi, Y., Nagata, N., & Inokuchi, S. (2004). The mental workload of a ships navigator using heart rate variability. *Interactive Technology & Smart Education*, *2*, 127–133.

Pauzie, A. (2008). A method to assess the driver mental workload: The driving activity load index (DALI). *IET Intelligent Transport Systems*, *2*, 315–322.

Posner, M. I., & Peterson, S. E. (1990). The attentional system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42.

Posner, M. I., & Rothbart, M. K. (1992). Attentional mechanisms and conscious experience. In A. D. Milner & M. D. Rugg (Eds.), *The neuropsychology of consciousness* (pp. 91–111). San Diego, CA: Academic Press.

Smith, M. E., Gevins, A., Brown, H., Karnik, A., & Du, R. (2001). Monitoring task load with multivariate EEG measures during complex forms of human-computer interaction. *Human Factors*, *43*, 366–380.

Sterman, M. B., Mann, C. A., Kaiser, D. A., & Suyenobu, B. Y. (1994). Multiband topographic analysis of a simulated visuomotor aviation task. *International Journal of Psychophysiology*, *16*, 49–56.

Strayer, D. L., Drews, F. A., & Crouch, D. J. (2006). A comparison of the cell phone driver and the drunk driver. *Human Factors*, *48*, 381–391.

Tsimhoni, O., Smith, D., & Green, P. (2004). Address entry while driving: Speech recognition versus a touch-screen keyboard. *Human Factors*, *46*, 600–610.

Veltman, J. A., & Gaillard, A. W. K. (1996). Physiological indices of workload in a simulated flight task. *Biological Psychology*, *42*, 323–342.

Walter, H., Vetter, S. C., Grothe, J., Wunderlich, A. P., Hahn, S., & Spitzer, M. (2001). The neural correlates of driving. *Neuroreport*, *12*, 1763–1767.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Sciences*, *3*, 159–177.

Wilson, G. F. (2001). In-flight psychophysiological monitoring. In F. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory monitoring* (pp. 435–454). Seattle, WA: Hogrefe and Huber.

Wilson, G. F. (2002). An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *International Journal of Aviation Psychology*, *12*, 3–18.

Wilson, G. F., & Russell, C.A. (2003). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors*, *45*, 635–643.

Wilson, G. F., & Russell, C. A. (2007).Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. *Human Factors*, *49*, 1005–1018.

Yamamoto, S., & Matsuoka, S. (1990). Topographic EEG study of visual display terminal VDT performance with special reference to frontal midline theta waves. *Brain Topography*, *2*, 257–267.

Shengguang Lei is working toward his PhD in human-machine systems in the Department of Psychology and Ergonomics at the Berlin Institute of Technology, Germany. He received his MSc in human-machine engineering in 2007 at Beijing Institute of Technology, China.

Matthias Roetting is a professor and the chair of human-machine systems in the Department of Psychology and Ergonomics at the Berlin Institute of Technology, Germany. He received his PhD in engineering in 2001 from Aachen University of Technology, Germany.

Date received: July 6, 2010
Date accepted: January 15, 2011