ASSESSING DRIVERS' VIGILANCE STATE DURING MONOTONOUS DRIVING

Eike A. Schmidt, Wilhelm E. Kincses, Michael Schrauf,[1]
Stefan Haufe, [2, 3] Ruth Schubert & Gabriel Curio[4]
[1]DaimlerChrysler AG, Group Research
[2]Technische Universität Berlin, Machine Learning group,
[3]Fraunhofer Institute FIRST, IDA,
[4]Neurophysics Group, Dept. of Neurology, Charité University-Medicine
Sindelfingen / Berlin, Germany
E-mail: eike_andreas.schmidt@daimlerchrysler.com

Summary: The differential effects of three hours of monotonous daytime driving on subjective (sleepiness, inattention, monotony), performance (choice reaction time), and physiological (EEG alpha power, P300-amplitude, heart rate) vigilance measures were examined. A linear degradation of drivers' subjective state, mean long reaction times (as opposed to short ones), P300-amplitude and parietal alpha power with time spent on the highway was identified. An improvement of the subjective measures towards the end of the driving task was not accompanied by any improvement in performance or physiological measures. This dissociation of self-assessment and objective vigilance measures has important implications for the design of modern driver assistant systems that aim to adapt to the driver's state.

INTRODUCTION

Vigilance describes "a state of readiness to detect and respond to certain specified small changes occurring at random time intervals in the environment" (Mackworth, 1957). The requirements of the driving task performed on long distance drives very much resembles this definition, because the ability to detect small changes in the environment is a crucial factor in safe driving. Accordingly, states of reduced vigilance, e.g., due to long and monotonous drives, are a main cause of traffic accidents (Thiffault & Bergeron, 2003). The development of driving assistance systems and active safety systems targeting low-vigilance conditions requires a comprehensive understanding and characterization of this cognitive state. For future applications, a reliable, real-time estimation of vigilance state and, more specifically, of the driver's reactivity, is desirable. The interdisciplinary project "FaSor" ("Fahrer als Sensor," which translates to "using the driver as a sensor") was conceived to contribute to this challenge by bringing together neurophysiologists, psychologists and computer scientists. Directly facing the hard classification challenge in real traffic situations, our goals were (1) to reliably obtain states of low vigilance by means of long monotonous highway driving, and (2) to investigate the time courses of measures reported in the literature to correlate with vigilance fluctuations. Special interest was given to the differential effects of driving duration on subjective, performance, and physiological measures. As an established measure of subjective sleepiness, we applied the Karolinska Sleepiness Scale (KSS; Åkerstedt & Gillberg, 1990). In order to acquire the subjective state more thoroughly, we constructed two further single-item scales regarding subjectively experienced inattention and monotony. To assess performance- and event-related physiological measures, a classical oddball paradigm was used. Long reaction times (as opposed to short ones), as well as amplitudes of the

P300 ERP components, were expected to be sensitive to changes in vigilance state (Williams et al., 1959; Koelega et al., 1992). Finally, frontal and parietal EEG alpha-power, as well as heart rate, were assessed; both are known to be sensitive to vigilance fluctuations (e.g., O'Hanlon & Kelly, 1977). As the current research was designed as a precursor to more sophisticated single subject analysis, in this study we wanted to validate the experimental setting in a group study and identify tonic changes in vigilance state. Therefore, we chose to evaluate blocks of 20 minutes' duration and to consider time as the only experimental factor.

METHODS

Procedure

Twenty-nine subjects (20 male, 9 female) with extensive driving experience drove 430 km (~267 miles) on a low-traffic German highway at a maximum speed of 130 km/h (~80 mph) using an upper class car (Mercedes Benz, S-Class). The course was situated on the A81 between Stuttgart and Singen, usually took about 3:30 h, and was performed during the day, between 13:00 and 17:00 hours, except for cases in which the experiment was terminated by the participants earlier. Three predefined turns were necessary and interrupted the continuous run at about 1:00, 1:40, and 2:20 h cumulated driving duration. To elicit the P300 evoked potential, the subjects had to perform an auditory oddball reaction time task during driving. The subjects responded to infrequent target tones by pressing a button fitted to their right thumb. The infrequent target tones (500 Hz, 20% probability) were presented in a random sequence mixed with frequent distractor tones (400 Hz, 80%), with an inter-stimulus interval (ISI) varying randomly between four and six seconds. It was ensured that the button could be easily pressed, no matter where the subject's hands were positioned on the steering wheel. Reaction times and response accuracy were continuously recorded. Every 20 minutes, single-item indicators of sleepiness, inattention and monotony regarding the last 20 minutes of driving were assessed verbally by the study investigator accompanying the experimental subject in the back seat throughout the drive. For all three scales, a low value (minimum: 1) indicated an extremely awake/attentive/varied state while a high value (maximum: 9) indicated an extremely sleepy/inattentive/monotonous state. EEG and electrocardiogram (ECG) were recorded from 128 electrodes (1000 Hz sampling rate, low cut-off: 0.016 Hz; high cut-off: 250 Hz). Unexpected and noticeable situations were logged by the investigator.

Performance Measures

To extract the mean long reaction times, for every block of 20 minutes, the mean of all reaction times above the 80%-percentile was calculated. Similarly, the average short reaction times (<20%-percentile) were identified for every block. The measure regarding the frequency of long reaction times corresponds to the number of reaction times per 20-minute block above the 80%-percentile of the whole session.

Physiological Measures

R-peaks were identified from the ECG using an automated algorithm, and average heart rate was calculated for every 20-minute block.

The EEG signal was sampled to 250 Hz, bandpass filtered (0.5 - 50 Hz) and channels with a standard deviation larger than 75 or smaller than 5 microvolts were excluded from further analysis. The same thresholds were applied to exclude artefactual time segments on a minute basis. Single-trial P300 ERP-components were extracted from the EEG raw signal using an individually fitted spatial filter that optimally separated the target-triggered P300 activity from that following distractor stimuli using linear discriminant analysis (LDA; see Parra et al., 2005). Every epoch was baseline-corrected with respect to the interval from -50 to +50 ms relative to stimulus onset, and the mean amplitude of the interval from +300 to +600 ms was extracted for further analysis. Single trial mean amplitudes were averaged for every 20-minute segment, which included about 40 target trials each.

In order to minimize ocular and muscular artefacts, independent component analysis (Jung, Makeig, Humphries, Lee, McKeown, Iragui & Sejnowski, 2000) was applied in a restrictive manner. Only those components carrying a temporal and spatial pattern resembling that of neural sources were accepted. The resulting EEG-signal was subjected to Fast Fourier Transformation (FFT) and the relative alpha (8 - 12 Hz) bandpower with respect to a 2 to 45 Hz reference band was calculated for electrode positions Fz (frontal) and Pz (parietal).

Data Reduction

As the goal of the experiment was to validly obtain monotony, data epochs that clearly lacked monotonous driving (i.e., communication between driver and investigator, turning points, workload-inducing driving situations, traffic jams and short stops) were discarded from analysis. Following a conservative criterion, the presence of one of the above factors led to the exclusion of one entire minute of data. Additional factors, such as missing EEG-data due to technical reasons, lack of compliance in two subjects and two fatigue-related break offs, led to a further reduction of sample size. Finally, 16 (12 male, 4 female; age: M = 29.8, SD = 8.4) complete data sets containing all measures were subjected to statistical analysis.

Statistical Analysis

A multivariate approach (MANOVA) was used for all within-subject comparisons to identify the effect of the factor time. All multivariate test criteria correspond to the same (exact) F-statistic. The level of α was set to .05 for all analysis. Whenever H_0 had to be rejected, the partial η^2 is reported as a measure of relative effect size. Significant results were subjected to post-hoc trend-analysis, applying polynomial contrasts. Linear and quadratic trends are reported.

RESULTS

Time courses of all measures are reported in Figure 1.

Subjective Measures

We successfully manipulated drivers' subjective state as shown by a considerable variation between mean minimum and maximum values for all three measures (see Table 1).

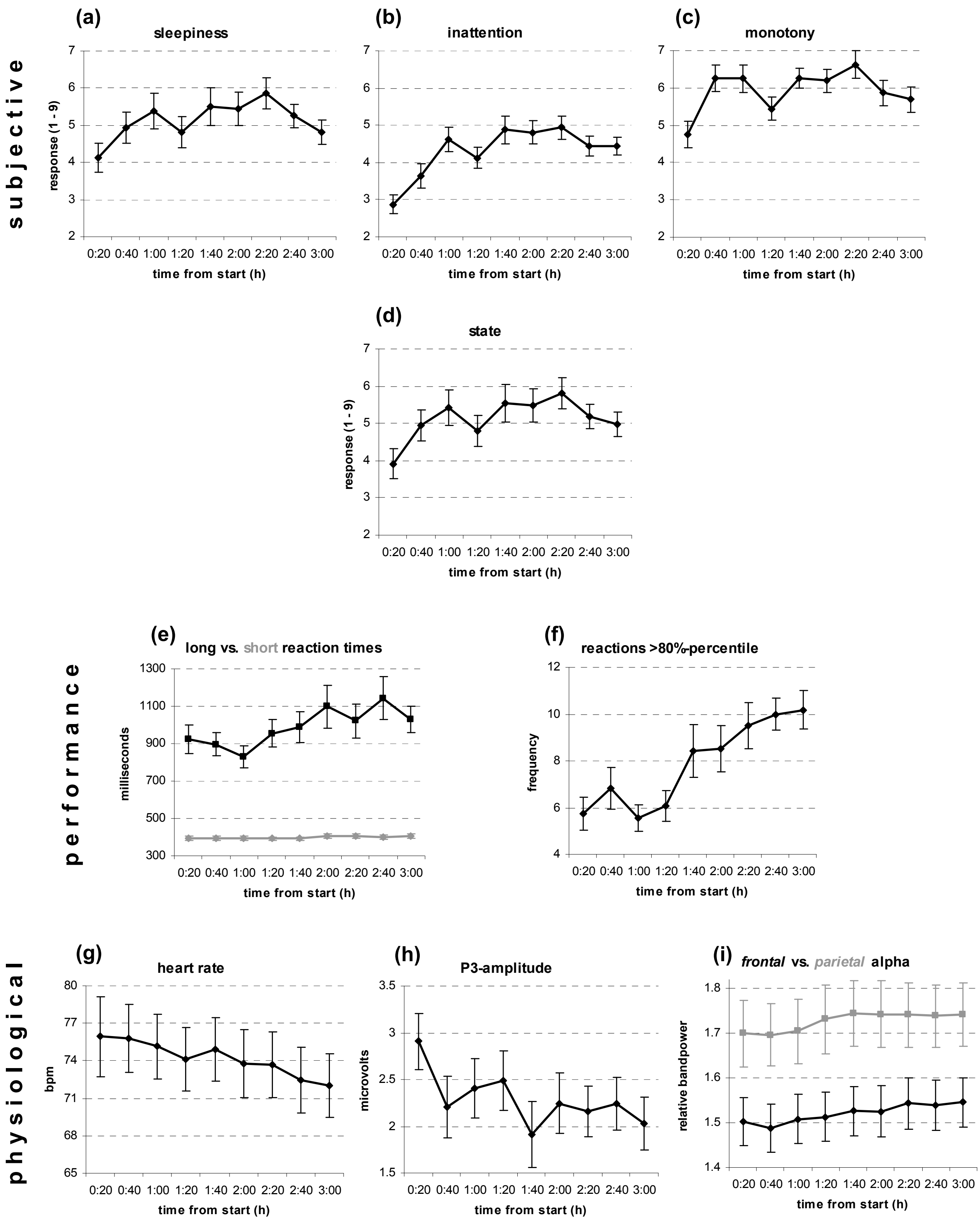

Figure 1. Dependent measures plotted versus time driven. Error bars indicate standard mean error.

Table 1. Descriptive statistics for subjective measures minimum, maximum and mean

		sleepiness	inattention	monotony
MIN	M	**3.63**	**2.75**	**3.94**
	SD	1.09	0.86	1.00
MAX	M	**6.69**	**5.63**	**7.56**
	SD	1.45	1.2	0.81
MEAN	M	**5.13**	**4.31**	**5.92**
	SD	1.36	0.9	0.65

Independent repeated measures MANOVAs revealed significant main effects of factors time and measure, while the interactions failed to reach significance (Table 2). In combination with high correlation coefficients between the time courses of the three measures, this implies a lack of variance in the temporal dynamics. For further analysis we therefore decided to collapse sleepiness, inattention and monotony into one conjunct measure of subjective state by calculating their mean for every incidence of assessment. A MANOVA testing the effect of the factor time on this subjective state measure revealed a significant effect ($F(8,8) = 11.68$; $p = .001$; $\eta^2 = .921$). Post-hoc trend analysis identified a significant linear ($F(1,15) = 7.14$; $p = .017$; $\eta^2 = .322$) as well as a more pronounced significant quadratic trend ($F(1,15) = 23.47$; $p < .001$; $\eta^2 = .610$).

Table 2. Results of subjective measures MANOVAs. Bold print indicates significant test statistic (α = .05). df(time) and df(time x measure): (8,8); df(measure): (1,15). In addition the Pearson correlation coefficient is reported for each combination of measures (df = 7).

	time			measure			time x measure			correlation	
	F	p	η^2	F	p	η^2	F	p	η^2	r	p
sleepiness x inattention	***11.31***	***.001***	***.919***	***14.10***	***.002***	***.485***	1.13	.432	.531	***.905***	***.001***
sleepiness x monotony	***6.56***	***.008***	***.868***	***5.19***	***.038***	***.257***	.569	.779	.363	***.924***	***<.001***
inattention x monotony	***7.81***	***.004***	***.886***	***36.77***	***<.001***	***.710***	1.635	.251	.621	***.777***	***.014***

Performance Measures

As shown in Table 3, there was no effect of time on short reaction times, while long reaction times, as well as the frequency of reaction times longer than the overall 80%-percentile, show a significant linear trend, whereas no quadratic trend is present.

Table 3. Results of performance measures MANOVAs. Bold print indicates significant test statistic (α = .05). df(time): (8,8); df(trends): (1,15).
Physiological Measures

	time			linear trend			quadratic trend		
	F	p	η^2	F	p	η^2	F	p	η^2
short RTs	.934	.537	.483						
long RTs	***5.05***	***.017***	***.835***	***6.58***	***.022***	***.305***	.02	.888	.001
freq. RTs > 80%	***3.98***	***.034***	***.799***	***30.45***	***<.001***	***.670***	.17	.678	.012

Significant effects of time were identified for P300-amplitude and parietal alpha power. Both measures show linear trends, whereas quadratic trends fail to reach significance.

Table 4. Results of physiological measures MANOVAs. Bold print indicates significant test statistic (α = .05). df(time): (8,8); df(trends): (1,15).

	time			linear trend			quadratic trend		
	F	p	η^2	F	p	η^2	F	p	η^2
heart rate	2.87	.078	.742						
p3-amplitude	***3.78***	***.039***	***.791***	***8.56***	***.010***	***.363***	4.20	.058	.219
frontal alpha	1.81	.210	.644						
parietal alpha	***4.98***	***.018***	***.833***	***6.63***	***.021***	***.306***	3.40	.085	.185

DISCUSSION

We replicated established findings by showing that long monotonous driving leads to a vigilance decrement over time (i.e., O'Hanlon & Kelly, 1977). In addition, the P300 was shown to be sensitive to vigilance variations in a real driving situation. To definitely disentangle the effects of driving duration from a simple habituation of the P300 (Polich, 1989), a comparative laboratory study might be necessary. Notably, the reaction time data support a relation between P300 decrement and vigilance reduction: the presence of a linear increase of mean long reaction times (while this is absent for mean short reaction times) corresponds well to classical findings (Williams et al., 1959) and can be explained by the increased number of attentional lapses. Auditory reaction time in a subsidiary task is a valid estimate for brake reaction time to unexpected obstacles, as shown by Laurell and Lisper in 1978. Therefore, it seems reasonable to infer from our data that after two hours of driving, the subjects' reactivity to unexpected traffic events was concurrently reduced.

The dissociation between the time course of drivers' subjective state (significant quadratic trend) and reaction time and physiology, respectively (no quadratic trends), over the last forty minutes suggests that even in a daytime drive of moderate duration, the ability to validly judge one's own personal state decreases with increasing time spent behind the wheel. The small-sized improved subjective state towards the end of the experiment might be due to the subjects' anticipated end of the monotonous ride. However this subjective internal recovery improved neither the physiological state nor the ability to perform well in the reaction time task. It seems interesting to discuss these findings in the light of work by Horne and Baulk (2004), who reported that their subjects were well aware of their physiological sleepiness over two hours of driving in a simulator. The dissociation between self-assessment and objective vigilance measures reported here bears important implications for the development of future adaptive driver assistance systems. If reactivity decreases significantly and the driver is not aware of it, he should either be informed about this misconception and/or potential warnings have to be presented earlier.

The mid- to long-term goal of our investigations is the real-time classification of vigilance with a high temporal resolution. This implies the engagement of machine learning algorithms that classify the respective cognitive state from EEG data as well as from other means. Most probably, we will engage a supervised real-time single trial-based machine learning algorithm, which means that it will be necessary to either identify a clear spatio-temporal EEG correlate of vigilance or to have an indirect vigilance indicative measure. However, the latter implies to provide sufficiently high temporal resolution in order to provide enough samples for the algorithm learning process. According to our findings, subjective measures are neither reliable enough nor do they have the temporal resolution to qualify as a valid labelling measure. According to our finding, reaction times are the most reliable measures for vigilance, since they also provide the highest significance for the driving task. However, the better the reaction times are embedded in driving task relevant secondary task, the more insight we expect to obtain regarding the state of vigilance.

REFERENCES

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience, 52*, 29-37.

Horne, J.A., & Baulk, S.D. (2004). Awareness of sleepiness when driving. *Psychophysiology, 41*, 161-165.

Jung, T.P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M.J., Iragui, V., & Sejnowski, T.J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology, 37*, 163-178.

Koelega, H.S., Verbaten, M.N., van Leeuwen, T.H., Kenemans, J.L., Kemner, C., & Sjouw, W. (1992). Time effects on event-related brain potentials and vigilance performance. *Biological Psychology, 34*, 59-86.

Laurell, H., & Lisper, H.-O. (1978). A validation of subsidiary reaction time against detection of roadside obstacles during prolonged driving. *Ergonomics, 21*, 81-88.

Mackworth, N.H. (1957). Vigilance. *The Advancement of Science, 53*, 389-393.

O'Hanlon J.F., & Kelly, G.R. (1977). Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation. *NATO Conf. Ser. 3*, 87-109.

Parra, L.C., Spence, C.D., Gerson, A.D., & Sajda, P. (2005). Recipes for the linear analysis of EEG. *Neuroimage, 28*, 326-341.

Polich, J. (1989). Habituation of P300 from auditory stimuli. *Psychobiology, 17*, 19-28.

Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis and Prevention, 35*, 381-391.

Williams, H.L., Lubin, A., & Goodnow, J.J. (1959). Impaired performance with acute sleep loss. *Psychological Monographs: General and Applied, 73(14)*, 1-26.