Transportation Research Part F 14 (2011) 251–260

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

The short-term effect of verbally assessing drivers' state on vigilance indices during monotonous daytime driving

Eike A. Schmidt [a,b,*], Michael Schrauf [c,d], Michael Simon [c,e], Axel Buchner [b], Wilhelm E. Kincses [c]

[a] Federal Highway Research Institute (BASt), Section U3: Traffic Psychology, Traffic Medicine, Brüderstraße 53, 51427 Bergisch Gladbach, Germany
[b] Heinrich-Heine-University Düsseldorf, Institut für Experimentelle Psychologie, Universitätsstr. 1, D-40225 Düsseldorf, Germany
[c] Daimler AG, Research and Development, GR/PAP, HPC 50-G024, D-71059 Sindelfingen, Germany
[d] University of Regensburg, Institute of Experimental Psychology, Universitätsstr. 31, 93053 Regensburg, Germany
[e] University of Tübingen, Wilhelm-Schickard-Institut für Informatik, Sand 13, D-72076 Tübingen, Germany

ARTICLE INFO

Article history:
Received 1 February 2010
Received in revised form 9 December 2010
Accepted 20 January 2011

Keywords:
Vigilance
Fatigue
Monotony
Driving
Self assessment
Fatigue countermeasures

ABSTRACT

To investigate the effects of verbal assessment of subjective driver state on objective indicators of vigilance state during a monotonous daytime drive, a real road driving study was conducted. During a 4-h drive participants' subjective state (sleepiness, inattention, monotony) was assessed every 20 min by an investigator accompanying the drive. The assessment procedure consisted of roughly 1 min of verbal interaction. Physiological indicators (EEG alpha spindle rate, blink duration, heart rate) revealed a significant improvement of vigilance state during the communication episode as compared to a pre-assessment baseline. The activation persisted for up to 2 min following the end of the verbal interaction. Reaction times supported these findings by indicating a significant decrease after the communication. The P3 amplitude of the auditory event-related potential did not show any consistent results. It can be concluded that a short verbal assessment has positive effects on drivers' vigilance state. However, these effects persist only for a very limited time. The implications of these findings for the frequency of verbal assessment during experimental studies and for the use of verbal communication as a fatigue countermeasure are discussed.

1. Introduction

The negative effects of prolonged monotonous night- and daytime driving on driver vigilance have been repeatedly investigated (e.g., Brookhuis & De Waard, 1993; Horne & Reyner, 1999; O'Hanlon & Kelly, 1977; Philip et al., 2005; Schmidt et al., 2009; Thiffault & Bergeron, 2003). In an attempt to reduce the number of fatigue-related accidents, numerous researchers have been working on the task of reliably detecting vigilance-related deficits in drivers (for reviews see Kecklund et al., 2006; Wright, Stone, Horberry, & Reed, 2007). The general absence of a single valid measure for every individual (Kecklund et al., 2006) makes it necessary to acquire a wide variety of vigilance indicators including subjective self-assessment measures of driver state. Subjective measures, often assessed verbally, bear the potential of influencing the state being under investigation (Kaida, Åkerstedt, Kecklund, Nilsson, & Axelsson, 2007) and by their nature cannot be recorded continuously. The objective of this study was to evaluate the magnitude and especially the duration of the potential influence of verbal assessment on driver state in real traffic. This should indicate how often a subjective assessment is feasible without significantly influencing the state under investigation. As the verbal assessment resembles a short communication between driver

---

* Corresponding author at: Federal Highway Research Institute (BASt), Section U3: Traffic Psychology, Traffic Medicine, Brüderstraße 53, 51427 Bergisch Gladbach, Germany Tel.: +49 176 2060 2075; fax: +49 211 811 5037.

E-mail addresses: eikeandreas.schmidt@uni-duesseldorf.de (E.A. Schmidt), michael.schrauf@daimler.com (M. Schrauf), michael.m.simon@daimler.com (M. Simon), axel.buchner@uni-duesseldorf.de (A. Buchner), wilhelm.kincses@daimler.com (W.E. Kincses).

1369-8478/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.trf.2011.01.005

and passenger, the presented findings also allow inferring its potential as a countermeasure to a vigilance decrement under monotonous daytime driving conditions.

1.1. Terminology: fatigue and sleepiness

Sleepiness and fatigue are often used synonymously. For clarification May and Baldwin (2009) introduced their model of fatigue, distinguishing between task-related active as well as passive fatigue and sleep-related fatigue, which most authors refer to as sleepiness or the difficulty in remaining awake (Philip et al., 2005). May and Baldwin state that the performance decrement – and therefore the task of detecting it – might be similar for all types of fatigue whereas countermeasures might only work for one or the other. Given that we investigated monotonous daytime driving with well-rested drivers, our focus is on task-related passive fatigue, although strictly speaking a minor influence of sleep-related fatigue could not be ruled out because our participants partly drove during their circadian afternoon dip.

1.2. Subjective assessment of driver state

Various single- and multi-item questionnaire measures of sleepiness have been developed and have been validated against both physiological as well as performance measures (e.g., Stanford Sleepiness Scale “SSS”: Hoddes, Zarcone, Smythe, Phillips, & Dement, 1973; Karolinska Sleepiness Scale “KSS”: Kaida et al., 2006; Åkerstedt & Gillberg, 1990). Apart from their limited objectivity, subjective questionnaire measures of sleepiness imply a tradeoff between the number of measurements in a given time period and the possibility of interfering with the process under investigation because answering questions may have activating effects and may thus reduce sleepiness. According to Kecklund et al. (2006) the measurement frequency of the KSS varies from once every 3–5 min down to just a few assessments per hour. Kaida et al. (2007) found that indeed the repeated rating of sleepiness reduced subjective post-test sleepiness, reduced EEG alpha power, and improved the subjective perception of performance. Interestingly, task performance did not improve. This led the authors to conclude that the effects of sleepiness were underestimated in the sleepiness ratings. Kaida et al. were interested in the long-term effects of sleepiness ratings for an entire experimental session. Therefore, conclusions about the immediate influence of the verbal sleepiness assessments were not obtained.

1.2.1. Indications from studies on countermeasures

Although little is known about the short-term effects of a subjective state assessment in a monotonous situation, related research on fatigue countermeasures might shed a light on this issue. A variety of experimental studies systematically evaluating the effects of potential countermeasures on sleep-related fatigue have been conducted by Reyner and Horne (1997, 1998). They concluded that the most effective and enduring countermeasure to sleepiness was a combination of caffeine intake and a short nap. In comparison, any positive effects of exposure to cold air and turning on the radio lasted only for about 15 min. However, questionnaire-based studies showed that in retrospective assessments between 25% and 35% of the drivers considered the engaging in a conversation with a passenger to be useful (Anund, Kecklund, Peters, & Åkerstedt, 2008; Maycock, 1997; Nordbakke & Sagberg, 2007). Therefore, at least some positive effects of investigator–participant communications are to be expected, if only for a brief post-communication interval.

1.2.2. Effect of conversations

To our knowledge the potentially activating effects of a verbal communication under monotonous driving conditions have been investigated in simulator studies only (Chan & Atchley, 2009; Drory, 1985; Gershon, Ronen, Oron-Gilad, & Shinar, 2009; Oron-Gilad, Ronen, & Shinar, 2008). Drory reported that long-haul truck drivers’ performance improved during a monotonous simulator drive when they were engaged in a short verbal task from time to time. Oron-Gilad et al. and Gershon et al. showed that a trivia game, carrying a verbal component, prevented deterioration of simulator driving performance and improved alertness assessed by physiological measures. This effect occurred immediately when engaged in the task but did not last longer than the task itself.

Despite this, studies on the distracting effects of cell phone and passenger conversations (Caird, Willness, Steel, & Scialfa, 2008; Drews, Pasupathi, & Strayer, 2008; Horrey & Wickens, 2006) have suggested that distraction might outweigh any positive effect of these types of conversations on driving performance. It is beyond the scope of this paper to disentangle these effects. We mainly focus on the effect *following* the communication.

1.3. Assessing drivers’ vigilance state

The vigilance measures applied in the research reported here can be classified as follows:

- Continuous physiological measures with high temporal resolution and the possibility to take measurements both during and after the investigator–driver communication.
- Discrete event-related measures elicited by an auditory stimulus (providing reaction times at the psychophysical, and event-related brain potentials at the neurophysiological side) that (a) need a sufficient number of stimuli for a reliable assessment and (b) cannot reasonably be recorded during the communication sequence.

1.3.1. Continuous measures

In accordance with our previous research we used the *alpha spindle rate*, a feature derived from the EEG alpha-band (7–13 Hz), as a sensitive EEG-based measure of a vigilance decrement induced by a monotonous driving situation (Simon et al., in press; Schmidt et al., 2009; see also Kecklund & Åkerstedt, 1993; Papadelis et al., 2007; Tietze & Hargutt, 2001). Apart from its high specificity to changes in vigilance, the main reason for preferring this measure to the classic EEG power measures is its robustness against external and internal noise and artifacts.

We also measured eye-blink duration because several authors reported *blink duration* to be their most informative occulomotoric parameter when measuring fatigue (Caffier, Erdmann, & Ullsperger, 2003; Ingre, Åkerstedt, Peters, Anund, & Kecklund, 2006). In a large sample study Schleicher, Galley, Briest, and Galley (2008) observed a gradual increase in blink duration with decreasing alertness.

Finally participants' *heart rate* was recorded as an indicator of the physical activation level, which has shown to be a sensitive indicator of vigilance changes in the driving context (O'Hanlon & Kelly, 1977).

1.3.2. Discrete event-related measures

In the context of a real driving situation, Laurell and Lisper (1978) demonstrated that an auditory secondary reaction time task was sensitive to changes in vigilance, and that it predicted brake reaction times. *Slow reactions* (as opposed to fast ones) seem to be particularly sensitive indicators of reduced vigilance (Graw, Kräuchi, Knoblauch, Wirz-Justice, & Cajochen, 2004; Williams, Lubin, & Goodnow, 1959) even in the driving context (Schmidt et al., 2009). Accordingly we implemented a simple secondary auditory task that most likely would not interfere with the motor requirements of the driving task, assuming that even minor reductions in vigilance should first be reflected in secondary task performance (subsidiary task paradigm: O'Donnell & Eggemeier, 1986). Participants were explicitly instructed to prioritize the primary task of driving and it was to be expected that the potentially high costs of major driving errors would also cause participants to give the highest possible priority to the driving task.

We further assessed the *amplitude* of the stimulus-induced *P3 event-related potential* (ERP, for a review see Polich, 2007) that has also been shown to be sensitive to changes in vigilance (Koelega et al., 1992; Schmidt et al., 2009) and can be interpreted as a measure of the processing depth of the auditory stimulus.

1.4. Hypothesis

We hypothesized that (1) the drivers' vigilance state should decrease with driving distance (main effect of distance). This should be reflected in an increase in subjective fatigue measures, alpha spindle rate, blink duration, and reaction times as well as a decrease in heart rate and P3 amplitude. We further hypothesized that (2) the verbal assessment of drivers' state by the investigator should improve drivers' vigilance state (main effect of communication). This should be reflected in a decrease in alpha spindle rate, blink duration, and reaction times and in an increase in heart rate and P3 amplitude.

2. Methods

2.1. Participants

Twenty-six right-handed participants (20 male, six female; age: $M = 26.6$, range: 21–40) with extensive driving experience (mean annual driving distance of approximately 20,500 km, i.e. 12,800 miles per year) were recruited on a voluntary basis for an "in-car EEG driving study". Participants were screened for a variety of exclusion criteria (handedness, auditory and visual disabilities, various illnesses), instructed to sleep regularly the night before the experiment, and to refrain from consuming caffeine in the morning on the day of the experiment. For their participation they received a monetary compensation of €100. The size of the sample available for data analysis was reduced by three subjects who aborted the drive due to heavy fatigue after less than 2:30 h driving duration which would have resulted in too few data points for a reliable analysis. As a result, 23 (18 male, five female; age: $M = 26.7$, range: 21–40) data sets containing all measures were available for statistical analysis.

2.2. Materials and procedure

Participants arrived at 9:00 am and signed an informed consent form. While the physiological recording equipment was applied, the participants completed the German versions of the morningness–eveningness–questionnaire (D-MEQ: Griefahn, Künemund, Bröde, & Mehnert, 2001) and the Edinburgh handedness inventory (Oldfield, 1971). Prior to the test-drive in real traffic the participants completed a monotonous simulated driving session of approximately 1.5 h duration starting at 10:00 am in a low-level fixed base simulator (Lane Change Task: Mattes, 2003). The results of the simulator study will be published elsewhere and are not subject of the present paper. To control for possible circadian (Folkard, 1997; Lenné, Triggs, & Redman, 1997) and nutritional effects (Smith & Miles, 1986) all participants had lunch at 11:30 am before a 30-min in-car EEG-baseline was recorded containing typical body movements (i.e. look over shoulder) and resting-EEG to gain experience on typical artifacts.

Following the German legal maximum uninterrupted driving duration for commercial drivers (4.5 h) the length of the drive was set to approximately 4 h and 15 min. To ensure as much monotony as possible, the participants drove on a low-traffic highway off rush hour times (A 81 Autobahn, south of Stuttgart, Germany). To minimize the systematic influence of route-specific factors, we decided to vary the exits where the participants were to change direction during the drive. This resulted in two different experimental courses of 475 km (about 297 miles; tuning points at 118, 190, 261 and 342 km) and 481 km (301 miles; tuning points at 176, 248, 303 and 358 km).

Participants started at 12:45 pm and returned, on average, after 4:08 h of driving, except for cases in which the experiment was terminated earlier by the participants or the investigator. The final sample included 19 subjects that completed the whole drive and four subjects that dropped out earlier but still supplied enough data for a reliable analysis (average break-off after 3:04 h of driving).

Participants had to be sufficiently rested upon arrival. They were informed that they could stop driving at any time without any monetary or other penalties. An investigator accompanied the participant in the car, continuously monitoring the driver and ready to intervene whenever necessary. The test car was a Mercedes Benz S-Class (W221). The participants' task was to drive at a speed not exceeding 130 km/h (approximately 80 mph; recommended maximum speed on German highways) and to comply with the traffic rules at all times. They were instructed to use automatic shift and to refrain from turning on the radio or using other in-car devices. Further, participants were asked not to talk to the investigator outside the communication episodes except for the case of expressing the wish to abort the driving. Participants were also instructed to avoid unnecessary movements in order to reduce artifacts in the EEG recording. The participants knew that they would be prompted for their subjective state frequently throughout the drive. They were not informed about when and how often they were to be asked by the investigator nor did they know about the researchers' particular interest in the effect of the communication episodes.

2.3. Communication episodes and subjective measures

In a prior research setting (Schmidt et al., 2009) the retrospective vigilance assessment occurred every 20 min. Despite this drivers experienced high levels of monotony and showed a pronounced vigilance decrement over time. We therefore decided to assess vigilance every 20 min. As subjective indicators we used the KSS (Åkerstedt & Gillberg, 1990) as a well-established single-item indicator of sleepiness and three similarly constructed items assessing inattention to the driving task, inattention to the auditory tones, and experienced monotony. The investigator verbally prompted the driver to judge his or her state concerning the four items on a nine point Likert-scale with regard to the previous 20 min of driving time. Each of these communication episodes lasted for approximately 1 min. The typical sequence of the dialogue containing the four subjective dimensions is displayed in Table 1.

2.4. Sections of analysis

In order to define a baseline for each participant that should reflect alert driving, we analyzed the first 20 min, that is, the section before the first communication. This baseline is descriptively reported in the figures for all measures and in case of the alpha spindle rate was used to derive an individual index. Of all communication episodes only those that were not corrupted by any turning point entered the analysis because performing a turn might have had an activating influence. The remaining episodes were divided by half into early and late communications. To keep the amount of data entering the analysis for the early and late conditions constant within each participant, in case of an odd number of episodes the middle episode did not enter the final analysis. In order to define a stable state prior to the communications for all measures an interval from $t = -5{:}30$ min to $t = -0{:}30$ min was evaluated relative to the start of the communication episode. Although very unlikely, we left a gap of 30 s prior to the communication episode to rule out any possible anticipation of the communication by the participants that could have affected their state prior to the communication. For the continuous measures the communication interval per se was analyzed. Finally, an interval of 5 min directly following the communication was analyzed. For the continuous measures this interval was further divided into five sections of 1-min length each. For every dependent measure and every interval the mean over all available episodes was calculated.

Table 1
Typical communication sequence.

Investigator	Participant
"The next state assessment is up."	*"O.K."*
"Concerning the time period since the last prompting, how would you describe your predominant state on a scale from 1-extremely alert- to 9-extremely sleepy, fighting sleep-?"	*"5"*
"... how attentively have you been driving on a scale from 1-extremely attentively- to 9-extremely inattentively-?"	*"4"*
"... how attentively have you been reacting to the tones on a scale from 1-extremely attentively- to 9-extremely inattentively-?"	*"4"*
"... and how did you perceive the drive on a scale from 1-extremely varied- to 9-extremely monotonous-?"	*"6"*

2.5. Performance measures

The participants were instructed that the auditory oddball reaction time task was only to be completed if they felt that it was safe to do so in a given driving situation and that they were free to halt their responses during the communication episodes. During the entire drive infrequent (oddball-) tones (500 Hz, 20% probability) were presented in a random sequence mixed with frequent tones of a lower pitch (400 Hz, 80%). The inter-stimulus interval varied randomly between 4 and 6 s. The constantly low pitch of the stimuli reduced their alerting potential to a minimum. In a 2-AFC paradigm the participants had to respond to the tones by pressing buttons fitted to their right or their left thumbs. They were informed about their initial assignment of tone and response hand right before the start of the experiment (lower-pitched tone: left thumb, higher-pitched tone: right hand; or vice versa). This assignment was reversed half way through the drive for each participant. The initial assignment was selected randomly with the goal that, at the end of the experiment, the two types of assignment would have occurred equally frequently. Responses faster than 200 ms and slower than 4000 ms were not analyzed. For every analysis epoch, and separately for each participant, the mean of all reaction times above the 80th-percentile was calculated as a measure of the participants' slow reactions, and the mean of all reaction times below the 20th-percentile as participants' fast reactions.

2.6. Physiological measures

Electroencephalogram (EEG), electrooculogram (EOG), and electrocardiogram (ECG) were recorded continuously throughout the entire drive from 64 electrodes (500 Hz sampling rate, low cut-off: 0.016 Hz; high cut-off: 250 Hz) using BrainAmp recording hard- and software (Brainproducts GmbH, Munich, Germany). A total of 62 of the 64 electrodes were cap-mounted according to the 10–20 system (61 EEG and 1 vEOG), one was applied below the right eye (vEOG) and one was applied to the participant's chest (ECG). Within this paper we only present the EEG data of Pz and CPz electrodes because alpha spindles and P3 are most prominent over these sites. Nevertheless, data of all electrodes were used for the process of artifact reduction and for classification efforts not presented here. Data were down-sampled to 128 Hz and low-pass filtered at 48 Hz. Artefactual channels were excluded from further analysis. In order to minimize the influence of muscle activity, eye blinks and technical noise, we de-noised the data using the extended infomax ICA algorithm (Lee, Girolami, & Sejnowski, 1999), available in the EEGLab toolbox (Delorme & Makeig, 2004). We used an automated algorithm to extract monochromatic bursts of activity within the alpha-band (7–13 Hz), which we call alpha spindles, and determined the amplitude, frequency, and duration of these activity bursts. In this paper we focus on the alpha spindle rate, defined as the number of alpha spindles per minute within each analysis epoch. Since the alpha rhythm is most prominent over parieto-occipital sites, we analyzed the signal of electrode Pz. To account for inter-individual differences an alpha spindle index was derived by dividing each section by the reference value of the first 20 min of driving.

To extract the P3 amplitudes the pre-processed and artifact-reduced EEG signal of electrode CPz was averaged time-locked to the presentation of the oddball stimulus. Trials with voltages ±100 μV were excluded. On average 34.3 trials (range 18–67) entered the analysis per interval. A baseline-correction (relative to −100 to 0 ms pre-stimulus time window) was applied. The P3 amplitude then was defined as the maximum value of the signal in a time window from 300 to 600 ms post-stimulus minus the average value of the baseline-epoch. The procedure was in accordance with the guidelines issued by Duncan et al. (2009).

In order to assess the duration of each eye blink in the vertical EOG, we identified its closing and reopening times in accordance with thresholds defined by Caffier et al. (2003) using an algorithm implemented in Matlab. Only blinks with durations between 50 and 500 ms and amplitudes larger than 100 μV entered the analysis. The mean duration over all blinks falling into the respective interval was calculated and entered into the statistical analysis.

R-peaks were identified from the ECG using an automated algorithm in Matlab. The average heart rate (beats per minute) was calculated for every interval.

2.7. Experimental design

The independent variables in our experimental design were the occurrence of the communication episodes during the drive (driving segment: early vs. late) and the time relative to the communication (communication: two sections for the event-related measures [−5:30 to −0:30 and +0:00 to +5:00 min]; seven sections for the continuous measures [−5:30 to −0:30 min, during communication, and five sections of 1 min length each following the communication]). Dependent variables were alpha spindle rate, heart rate, blink duration, slow and fast reactions, and P3 amplitude.

An a priori statistical power analysis using G∗Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007) showed that in order to detect effects of $f = .50$ of the communication within-subject variable assuming a population correlation among the levels of this variable of $\rho = .50$ (estimated from pilot data) and given $\alpha = \beta = .05$, between $N = 14$ and $N = 16$ participants were needed (depending on the number of levels of the communication variable; see previous paragraph). A sensitivity power analysis showed that given a final sample of $N = 23$, effects of size $f = .30$ to $f = .39$ could be detected under otherwise identical conditions. A multivariate approach (MANOVA) was used for all within-subject comparisons to identify the effects of the driving segment and communication factors for each dependent measure. All multivariate test criteria correspond to the same (exact) F-statistic, which is reported. The level of α was set to .05 for all analysis. Partial η^2 is reported as a measure

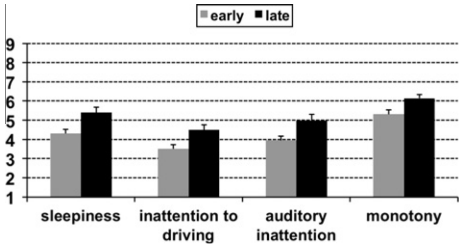

Fig. 1. Subjective measures for early vs. late communication episodes.

Table 2
Statistical results for continuous physiological measures.

	Driving segment			Communication			Driv. segm. × communication		
	$F(1, 22)$	p	η^2	$F(6, 17)$	p	η^2	$F(6, 17)$	p	η^2
Alpha spindle rate	15.84	<.001	.419	7.07	<.001	.714	1.02	.447	n.s.
Blink duration	4.07	.056	n.s.	6.14	.001	.684	.71	.684	n.s.
Heart rate	20.80	<.001	.486	10.68	<.001	.790	1.25	.329	n.s.

of relative effect size. Statistically significant results of the continuous measures were subjected to post hoc contrast analysis using the PRE-section as the fixed reference. In these cases the level of α was Bonferroni-Holm corrected (Holm, 1979) so as to avoid alpha error accumulation. For clarity of presentation only significant contrasts are reported.

3. Results

It was confirmed that all participants were right-handed (all handedness-indices >0). The D-MEQ results showed that there were no participants falling into any of the extreme chronotype groups (moderate evening-type: 7; neutral type: 13; moderate morning-type: 3).

On average about eight communication episodes entered the analysis for each participant (range: 4–12). The mean starting time of the early episodes was 2 h before the mean starting time of the late ones. Comparison of early and late communication episodes revealed that the late episodes were significantly shorter (62 s (early) vs. 50 s (late); paired sample t-test: $t(22) = 5.19$; $p < .001$; $\eta^2 = .550$). For the subjective measures participants reported significantly higher judgments of all measures of vigilance for late as compared to early communication episodes (Fig. 1; sleepiness: $t(22) = 4.20$; $p < .001$; $\eta^2 = .445$; inattention to driving: $t(22) = 4.37$; $p < .001$; $\eta^2 = .465$; auditory inattention: $t(22) = 3.78$; $p = .001$; $\eta^2 = .394$; monotony: $t(22) = 3.86$; $p = .001$; $\eta^2 = .404$).

3.1. Continuous physiological measures

The results of the MANOVAs testing the effects of driving segment and communication on the continuous physiological measures are summarized in Table 2. Both variables exerted a significant influence on alpha spindle rate and heart rate. Alpha spindle rate (Fig. 2) was significantly higher around late as compared to early episodes. Relative to the pre-communication baseline, a significant reduction of alpha spindle rate was observed during the communication episodes ($F(1, 22) = 24.59$; $p < .001$; $\eta^2 = .528$). The reduction persisted for the subsequent minute ($F(1, 22) = 18.24$; $p < .001$;

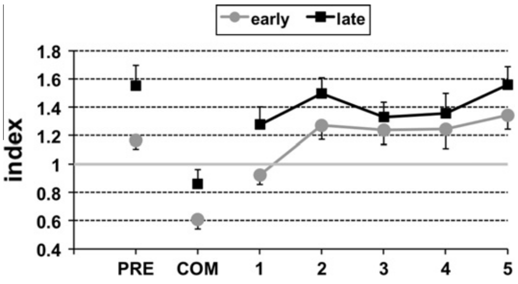

Fig. 2. Alpha spindle rate prior to (PRE), during (COM), and in the minutes following the communication episode (1–5) for early and late driving segments. Error bars represent the standard errors of the means. The grey horizontal line represents the baseline level during the first 20 min of driving.

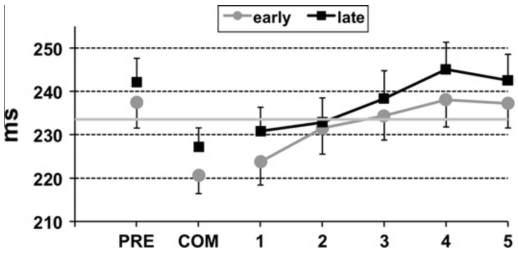

Fig. 3. Blink duration prior to (PRE), during (COM), and in the minutes following the communication episode (1–5) for early and late driving segments. Error bars represent the standard errors of the means. The grey horizontal line represents the baseline level during the first 20 min of driving.

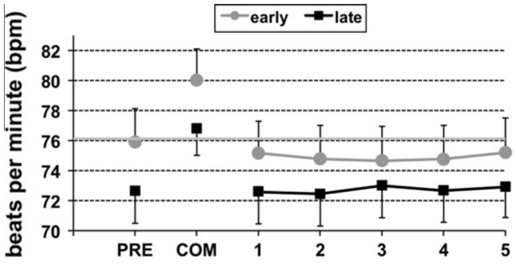

Fig. 4. Heart rate prior to (PRE), during (COM), and in the minutes following the communication episode (1–5) for early and late driving segments. Error bars represent the standard errors of the means. The grey horizontal line represents the baseline level during the first 20 min of driving.

$\eta^2 = .453$). The effect of driving segment on blink duration (Fig. 3) just fell short of the preset criterion of statistical significance ($F(1,22) = 4.07$; $p = .056$). The average blink duration was significantly reduced during the communication ($F(1,22) = 17.99$; $p < .001$; $\eta^2 = .450$) relative to the pre-communication baseline. This effect persisted for two post-communication minutes (first minute: $F(1,22) = 17.46$; $p < .001$; $\eta^2 = .442$; second minute: $F(1,22) = 12.85$; $p = .002$; $\eta^2 = .369$). Heart rate (Fig. 4) was significantly reduced around late as compared to early episodes. Relative to the pre-communication baseline, heart rate increased significantly during the communication ($F(1,22) = 44.86$; $p < .001$; $\eta^2 = .671$), but returned to the pre-communication level during the first post-communication minute. In the second post-communication minute, heart rate fell below the pre-communication level ($F(1,22) = 8.88$; $p = .007$; $\eta^2 = .287$).

3.2. Discrete event-related measures

The results of the MANOVAs testing the effects of driving segment and communication on the discrete stimulus based measures are summarized in Table 3. As expected, we observed neither an effect of driving segment nor an interaction of driving segment and communication on the mean of the fast reaction times (Fig. 5a). There was, however, a significant effect of the communication variable reflected in a decrease in the mean of the fast reaction times following the communication. In contrast, for the mean of the slow reaction times (Fig. 5b) we observed significant main effects for both driving segment and communication with an increase in slow reaction times in later sections of the drive and a decrease in slow reaction times following a communication episode. There were no significant effects on the P3 amplitude (Fig. 6).

4. Discussion

The indices of subjective driver state as well as the objective vigilance measures – that is, alpha spindle rate, heart rate, and slow reaction times – all showed a significant effect of driving segment. This pattern replicates findings from an earlier study (Schmidt et al., 2009). Blink duration also showed a tendency towards an increased average duration with increasing

Table 3
Statistical results for discrete event-related measures.

	Driving segment			Communication			Driv. segm. × communication		
	$F(1,22)$	p	η^2	$F(1,22)$	p	η^2	$F(1,22)$	p	η^2
Fast reactions	3.03	.096	n.s.	10.16	<.001	.316	.39	.541	n.s.
Slow reactions	15.27	<.001	.410	5.65	.027	.204	.53	.473	n.s.
P3 amplitude	.00	.975	n.s.	.99	.330	n.s.	3.29	.084	n.s.

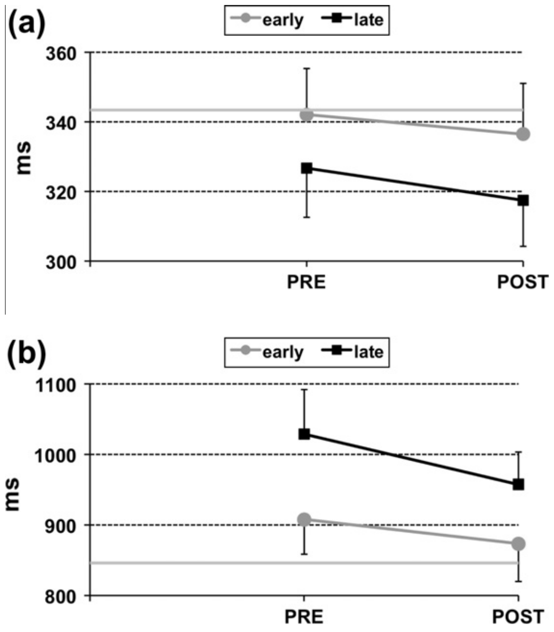

Fig. 5. Means of fast (a) and slow (b) reaction times prior to (PRE) and following (POST) the communication episode for early and late driving segments. Error bars represent the standard errors of the means. The grey horizontal lines represent the baseline levels during the first 20 min of driving.

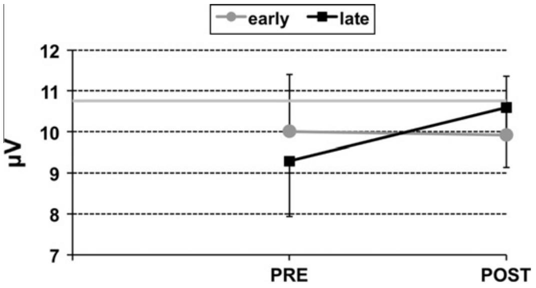

Fig. 6. P3 amplitude prior to (PRE) and following (POST) the communication episode for early and late driving segments. Error bars represent the standard errors of the means. The grey horizontal line represents the baseline level during the first 20 min of driving.

distance driven. It may be concluded that monotonous driving resulted in a vigilance decrement in the present study and that therefore the applied methodology was successful in experimentally inducing the intended effects on drivers' state. We suspect the absence of this effect for the P3 amplitude (which was a sensitive measure of fatigue in the Schmidt et al., 2009 study) to be based on an insufficient number of trials entering the averaging procedure for each condition resulting in a poor signal-to-noise ratio.

During the communication episode all continuous measures indicated a clear improvement of vigilance in terms of a decrease in alpha spindle rate and blink duration and an increase in heart rate. This pattern suggests that a significant activation of the participants was induced by the communication. As mentioned in the introduction, it is not possible to infer, from the present data, the net effect of communication-induced driver activation on the positive side and distraction on the negative side. A reasonable assessment of this issue will likely require extended simulator studies in which critical variables such as brake reaction times can be studied in sufficient detail.

Concerning the duration of the effect elicited by the communication *following* the interaction, the continuous physiological measures lead to consistent results. Both alpha spindle rate and blink duration indicated an activating effect persisting after the end of the communication episode which, however, was limited to just 1 min (alpha spindle rate) or 2 min (blink duration). Heart rate, in contrast, did not show any persisting effect and even moved below the pre-communication level after the communication. Thus, we conclude that the positive effects of verbal assessments on drivers' vigilance states may last beyond the communication episodes, but they do so only for a very short period of time.

Given this, we further conclude that an assessment interval of 5 min – the lower boundary recommended by Kecklund et al. (2006) – seems very reasonable if the goal is not to affect driving-induced fatigue. Strictly speaking, however, we cannot draw firm conclusions about possible cumulative effects of the repeated assessment (Kaida et al., 2007). This would have required running a control condition without any verbal assessments.

Unfortunately due to its event-related nature the reaction time data did not allow us to reliably estimate the duration of the effect at a sufficient temporal resolution. Still, it is clear from the data that following the communication the means of slow as well as fast reactions decreased significantly. The effect for the fast reactions was rather surprising, because prior research had shown fast reactions to be insensitive to changes in vigilance (e.g., Schmidt et al., 2009). It may be that the interaction with the investigator induces self-awareness and as such draws the driver's attention to the fact that there is a reaction time task to be accomplished while driving. As a result, greater priority may be given to the secondary task, as a result of which even the fast reactions become faster.

A methodological issue inherent in the present study lies in the fact that, trivially, the post-communication epochs followed the pre-communication epochs. Therefore driving segment and communication were to some extent confounded, which could have resulted in a small underestimation of the activating effect of the communication episodes because vigilance decreased as a function of the distance driven. However, we believe that the amount by which the short-term activation was reduced by the underlying long-term fatigue is negligible.

Due to the fact that the communication process comprises various stages of cognitive and motor actions, future research is needed in order to analyze in more detail the factors that cause the activating effect. Candidates are the social interaction with another human being, the self-awareness induced by the requirement to think about ones own state, or the mere effect of talking or listening. Also, the degree to which the present results generalize to an everyday conversation while driving remains unresolved. For instance, on the one hand our participants only answered fairly routine questions, which may not be too activating per se. On the other hand, answering these questions may have been particularly activating due to their generating a certain degree of self-awareness. In real-world situations, the topics of the conversation may be much more variable (and likely more interesting) while at the same time being less self-referential. Finally, it is probably fair to say that considering in-car communications only in terms of the risks they imply without considering their positive effects – especially under conditions of monotonous driving conditions – is not appropriate, although the positive effects seem to be rather limited.

5. Conclusion

Verbal assessment of driver state causes a direct improvement of driver vigilance state as indicated by alpha spindle rate, blink duration, and heart rate. The post-conversation temporal extension of this positive effect seems to be rather limited, as is indicated by the fact that the physiological vigilance measures returned to their pre-communication levels after 2 min at the latest. From these findings we conclude (a) that verbal assessments of fatigue are feasible even at a rate of one every 5 min, and (b) that there are clear positive effects of a verbal communication as a countermeasure for fatigue during the communication per se, but these positive effects can no longer be measured 2 min after the communication.

Acknowledgements

This research was supported by BMBF (German Federal Ministry of Education and Research) Grant 16SV2233. We thank Prof. Gabriel Curio, Dr. Ruth Schubert and Dr. Martin Fritzsche for valuable discussions on the experimental design; Andreas Pröttel and Sven Willmann for support with data recording and analysis; Michael Sachse and Ying Lee for their help with data collection.

References

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

Anund, A., Kecklund, G., Peters, B., & Åkerstedt, T. (2008). Driver sleepiness and individual differences in preferences for countermeasures. *Journal of Sleep Research, 17*, 16–22.

Brookhuis, K., & De Waard, D. (1993). The use of psychophysiology to assess driver status. *Ergonomics, 36*(9), 1099–1110.

Caffier, P. P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology, 89*, 319–325.

Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis and Prevention, 40*, 1282–1293.

Chan, M., & Atchley, P. (2009). Effects of cell phone conversations on driver performance while driving under highway monotony. In *Proceedings of the fifth international driving symposium on human factors in driver assessment, training and vehicle design, Big Sky, MT* (pp. 140–146).

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.

Drews, F. A., Pasupathi, M., & Strayer, D. L. (2008). Passenger and cell phone conversations in simulated driving. *Journal of Experimental Psychology: Applied, 14*(4), 392–400.

Drory, A. (1985). Effects of rest and secondary task on simulated truck-driving task performance. *Human Factors, 27*(2), 201–207.

Duncan, C., Barry, R., Connolly, J., Fischer, C., Michie, P., Näätänen, R., et al (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology, 120*(11), 1883–1908.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Folkard, S. (1997). Black times: Temporal determinants of transport safety. *Accident Analysis and Prevention, 29*(4), 417–430.

Gershon, P., Ronen, A., Oron-Gilad, T., & Shinar, D. (2009). The effects of an interactive cognitive task (ICT) in suppressing fatigue symptoms in driving. *Transportation Research Part F, 12*, 21–28.

Graw, P., Kräuchi, K., Knoblauch, V., Wirz-Justice, A., & Cajochen, C. (2004). Circadian and wake-dependent modulation of fastest and slowest reaction times during the psychomotor vigilance task. *Physiology and Behavior, 80*, 695–701.

Griefahn, B., Künemund, C., Bröde, P., & Mehnert, P. (2001). Zur Validität der deutschen Übersetzung des morningness–eveningness–questionnaires von Horne und Östberg. *Somnologie, 5*, 71–80.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of sleepiness: A new approach. *Psychophysiology, 10*, 431–436.

Holm, S. (1979). A simple sequentially rejective multiple teste procedure. *Scandinavian Journal of Statistics, 6*, 65–70.

Horne, J., & Reyner, L. (1999). Vehicle accidents related to sleep: A review. *Occupational and Environmental Medicine, 56*, 289–294.

Horrey, W. J., & Wickens, C. D. (2006). Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors, 48*, 196–205.

Ingre, M., Åkerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2006). Subjective sleepiness, simulated driving performance and blink duration: Examining individual differences. *Journal of Sleep Research, 15*, 47–53.

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., et al (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology, 117*, 1574–1581.

Kaida, K., Åkerstedt, T., Kecklund, G., Nilsson, J. P., & Axelsson, J. (2007). The effects of asking for verbal ratings of sleepiness on sleepiness and its masking effects on performance. *Clinical Neurophysiology, 118*, 1324–1331.

Kecklund, G., & Åkerstedt, T. (1993). Sleepiness in long distance truck driving: An ambulatory EEG study of night driving. *Ergonomics, 36*, 1007–1017.

Kecklund, G., Åkerstedt, T., Sandberg, D., Wahde, M., Dukic, T., Anund, A., & Hjälmdahl, M. (2006). *State of the art review of driver sleepiness. DROWSI project report, deliverable 1.1*. <www.vti.se/11934.epibrw> Accessed 03.01.10.

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.

Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and super-Gaussian sources. *Neural Computation, 11*, 417–441.

Lenné, M. G., Triggs, T. J., & Redman, J. R. (1997). Time of day variations in driving performance. *Accident Analysis and Prevention, 29*(4), 431–437.

Mattes, S. (2003). The lane-change-task as a tool for driver distraction evaluation. In H. Strasser, K. Kluth, H. Rausch, & H. Bubb (Eds.), *Quality of work and products in enterprises of the future* (pp. 57–60). Stuttgart: Ergonomia.

May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasures technologies. *Transportation Research Part F, 12*, 218–224.

Maycock, G. (1997). Sleepiness and driving: The experience of UK car drivers. *Accident Analysis and Prevention, 29*, 453–462.

Nordbakke, S., & Sagberg, F. (2007). Sleepy at the wheel: Knowledge, symptoms and behaviour among car drivers. *Transportation Research Part F, 10*, 1–10.

O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.). *Handbook of perception and human performance. Cognitive processes and performance* (Vol. II, pp. 42/1–42/49). New York: Wiley.

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. In R. Mackie (Ed.), *Vigilance* (pp. 87–109). New York: Plenum Press.

Oldfield, R. C. (1971). The assessment and analysis of handedness. The Edinburgh inventory. *Neuropsychologia, 9*, 97–113.

Oron-Gilad, T., Ronen, A., & Shinar, D. (2008). Alertness maintaining tasks while driving. *Accident Analysis and Prevention, 40*, 851–860.

Papadelis, C., Chen, Z., Kourtidou-Papadeli, C., Bamidis, P. D., Chouvarda, I., Bekiaris, E., et al (2007). Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. *Clinical Neurophysiology, 118*, 1906–1922.

Philip, P., Sagaspe, P., Taillard, J., Valtat, C., Moore, N., Åkerstedt, T., et al (2005). Fatigue, sleepiness, and performance in simulated versus real driving conditions. *Sleep, 28*(12), 1511–1516.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10), 2128–2148.

Reyner, L. A., & Horne, J. A. (1997). Suppression of sleepiness in drivers: Combination of caffeine with a short nap. *Psychophysiology, 34*, 721–725.

Reyner, L. A., & Horne, J. A. (1998). Evaluation of "in car" countermeasures to driver sleepiness: Cold air and radio. *Sleep, 21*, 46–50.

Schleicher, R., Galley, N., Briest, S., & Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings: Looking tired? *Ergonomics, 51*(7), 982–1010.

Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accident Analysis and Prevention, 41*, 1087–1093.

Simon, M., Schmidt, E. A., Kincses, W. E., Fritzsche, M., Bruns, A., Aufmuth, C., et al. (in press). EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. *Clinical Neurophysiology*, doi:10.1016/j.clinph.2010.10.044.

Smith, A. P., & Miles, C. (1986). The effects of lunch on cognitive vigilance tasks. *Ergonomics, 29*(10), 1251–1261.

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

Tietze, H., & Hargutt, V. (2001). *Zweidimensionale analyse zur Beurteilung des Verlaufs von Ermüdung. 43. Tagung experimentell arbeitender Psychologen, Regensburg, GER*. <http://www.psychologie.uni-wuerzburg.de/methoden/texte/2001_tietze_hargutt_Zweidimensionale_Analyse.pdf> Accessed 19.01.10.

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

Wright, N. A., Stone, B. M., Horberry, T. J., & Reed, N. (2007). *A review of in-vehicle sleepiness detection devices. Published project report 157. TRL Limited, UK, Crowthorne*. <http://www.masstransitmag.com/pdf/2007/mar/ABMT328Sleepy.pdf> Accessed 19.01.10.