Blinks and saccades as indicators of fatigue in sleepiness warners: looking tired?

R. Schleicher[1], N. Galley, S. Briest and L.Galley

Robert Schleicher
Deutsche Telekom Laboratories, Berlin University of Technology
Ernst-Reuter-Platz 7, 10587 Berlin, Germany.

Niels Galley
Department of Anatomy, University of Cologne
Joseph-Stelzmann-Strasse 9, 50931 Cologne, Germany

Susanne Briest
German Aerospace Center, Institute of Transportation Systems
Lilienthalplatz 7, 38108 Braunschweig, Germany

Lars Galley,
DaimlerChrysler AG, Customer Research Center
Leibnizstr. 2, 71032 Böblingen, Germany

[1] Corresponding author. Email: robert.schleicher@tu-berlin.de

DOI : 10.1080/00140130701817062

The present study examines changes in a variety of oculomotoric variables as a function of increasing sleepiness in 129 participants, who have been passed through a broad range of subjective alertness. Up to now, spontaneous eyeblinks are the most promising biosignal for in-car sleepiness warners. Reviewing the current literature on eye movements and fatigue, we provide experimental data including additional indicative oculomotoric parameters; we also assessed interindividual differences in our own experiments. Here, self-rated alertness decreased over six steps on average and proved itself a reliable measurement. Regarding oculomotoric parameters, blink duration, delay of lid reopening, blink interval, and standardised lid closure speed were identified as the best indicators of subjective as well as objective sleepiness. Saccadic parameters and fixation durations also showed specific changes with increasing sleepiness. Substantial interindividual differences in all of these variables were illustrated. Oculomotoric parameters were linked to three different components of sleepiness while driving, a) deactivation, b) decreasing attention, resulting in disinhibition of spontaneous blinks and reflexive saccades and c) increasing attempts of self-activation. Finally, implications for the development of drowsiness detection devices were discussed.
Keywords: fatigue, subjective rating scales, blink, saccade, fixation, EOG.

1. Introduction

In the last couple of years, camera-based drowsiness detection devices using lid cleft and lid-movement parameters have undergone intensive development (Hargutt, 2003; Svensson, 2004; Thorslund, 2003; Wierwille, 1999). Despite initial results that appear promising, many interesting ideas do not seem to get beyond prototypes (Hagenmeyer et al., 2007), partly due to large interindividual differences. Based on our own experiences in this field, we would like to describe some of the typical difficulties and provide data that show alternative ways out of these dilemmas.

In general, the development of a drowsiness detector requires the following components:

1. a reliable measure of the subject's current drowsiness.
2. a well-defined threshold for sleepiness at which the ability to perform is substantially reduced.
3. objective behavioural symptoms and/or biosignals (including appropriate sensors and software) that indicate such a critical point of sleepiness *in advance*,
4. a body of knowledge that can be applied on an *individual* level. As large variations in biosignal changes related to fatigue are reported to occur in an individual (Ingre, Akerstedt, Peters, Anund, & Kecklund, 2006; Van Dongen et al., 2004; Van Dongen et al., 2005), group means are insufficient. One has to examine to what extent a possible indicator or threshold reflects overtiredness in each individual subject and then it is necessary to show how deviant cases can still be identified properly.

The first step in developing a drowsiness detector based on physiological parameters is to define an independent measure of fatigue that can then be used to evaluate changes in the signals of interest (i.e. lid movements) (Lal & Craig, 2001a).

1.1 Measuring fatigue

1.1.1 The EEG. The most important signal in sleep research is spontaneous EEG. Unfortunately, it is hard to measure and analyse under field conditions. Additionally, EEG-based sleep research has had great difficulties in specifying sleep *onset*: Is onset a sudden or gradual change? It may even reflect oscillation. Moreover, its characteristic parameters are not undisputed in detail (for a review see Ogilvie *et al.*1989). In some studies, the disappearance of Alpha-activity in favour of Theta and/or Delta-activity is viewed as crucial (Cajochen et al., 1995; Gevins et al., 1995; Wright et al., 1995), while in other studies the important event is an increase of power in the broad Alpha-Theta-band (Akerstedt & Folkard, 1994;

Cajochen et al., 1995; Horne & Baulk, 2004). Yet other define sleep onset as the onset of sleep stage 2 and the appearance of its typical phenomena like sleep spindles and K-complexes (Hasan & Broughton, 1994).

Some authors prefer a broad diagnosis that includes behavioural data (Webb, 1994), such as the stopping of saccades and blinks that normally not only occur with closed eyes but can also be found in light sleep stage 1. However, in sleep laboratories eye movements are usually recorded with the EOG with an insufficient sample rate. Moreover, electrodes are placed in such a way that clear differentiation of saccades, slow eye movements, and blinks is impossible (i.e. 'diagonally' instead of horizontal and vertical channels strictly separated).

1.1.2 Performance measures. Performance measures such as reaction times as indicators of sleepiness have traditionally been appraised in different ways and this is still the case today. On the one hand, changes in performance level as represented by increased reaction times are declared to be a defining criterion of sleep onset: *'Functional sleep onset is, by definition performance based'* (Kribbs & Dinges 1994, p.114). Advocates of performance detoriations as the most convincing objective indicator of sleepiness (Balkin et al., 2004; Belenky et al., 2003; Dinges & Kribbs, 1991; Kribbs & Dinges, 1994) frequently use the psychomotoric vigilance test (PVT). This test is really troubled by the fact that it interrupts any ongoing activity in the order of up to twenty minutes, during which the user has to observe a small display constantly and stop a LED-counter that starts running randomly – a task that is clearly quite monotonous and itself may well *induce* drowsiness or lapses of attention.

On the other hand, several studies found intermediate reaction-time tasks to be insensitive to fatigue (Baulk et al., 2001; Gillberg et al., 1996). Similarly, Milosevic (1997) reported only a non-significant increase in acoustic reaction time from 211 to 222 ms after 7.5 hours of driving, whereas visual reaction time at least increased significantly from 234 to 261 ms. Porcu *et al.* (1998) observed that simple, less time-dependent visuomotoric tasks remained unaffected during a night of sleep deprivation; only performance in more complex tasks showed deterioration. For applied settings, the question arises, what level of performance change might be an appropriate criterium: For example, is continuing slight inaccuracies in lane keeping sufficient, or is it necessary to have the occurrence of major mishaps? As noted repeatedly, performance can be stabilized against deteriorating effects of fatigue by mobilizing additional mental resources to fulfil the required task (Galley, 1998; Hockey, 1993; Matthews & Desmond, 2002); therefore, error patterns may vary with different levels of fatigue (Dorrian et al., 2007) as well as the subject's age (Campagne et al., 2004; Philip, Taillard et al., 2003) or gender (Blatter et al., 2006). Consequently, it may be difficult to come up with general performance measures indicative of driver fatigue. As Thiffault and Bergeron (2003, p. 387) summarize: *'In other words, even if studies concur to say that these measures are accurate, the use of steering behavior to study driver fatigue remains exploratory.'*

1.1.3 Subjective measures. Next to the gold standard EEG and objective indicators like changes in performance or behavioural signs (e.g. yawning) sleepiness can also be evaluated by subjective measures using self-assessment, such as *self-rated alertness*. A widespread 9-step subjective scale is the Karolinska Sleepiness Scale (KSS). In its modified version by Horne & Reyner (1995), it contains the following descriptions: 1= Extremely alert, 2 = Very alert, 3 = Alert, 4= Rather alert, 5 = Neither alert nor sleepy, 6 = Some signs of sleepiness, 7 = Sleepy, but no effort to keep alert, 8 = Sleepy, *some effort to keep alert*, 9 = Very sleepy, *great effort to keep alert, fighting sleep.*

The expressions *some effort* and *great effort* in stages 8 and 9 point to the fact that occasional falling asleep might be more than just a decrease of general arousal: it may involve an attempt to stay awake with the consequence of a 'paradoxical' increase in activation. In this case, falling asleep is something completely different than simply giving in to the powerful urge to rest (Lavie, 1991, 2001). This then constitutes a serious dilemma, because the instructions in laboratory studies or possible live-threatening consequences in real life demand at least a rudimentary form of wakefulness. Under such circumstances,

one can observe considerable individual differences as well as paradoxical physiological changes (Van Dongen et al., 2004; Van Dongen et al., 2005).
Given its susceptibility to manipulation, subjective sleepiness was traditionally held in low regard and many researchers even refrain from collecting it in their experiments. However, Horne & Baulk (2004) did find a good correspondence of EEG-based indicators of sleepiness and subjective ratings. Similarly, Ingre et al. (2006a, 2006b) report a close relation of self-rated fatigue with driving errors and blink duration, although the correspondence of subjective and objective measures is better in simulator as compared to real-life driving (Philip, Sagaspe et al., 2003; Philip et al., 2005).

It is a well established fact that in laboratory conditions that are monotonous without live-threatening consequences, there is a rapid rise in fatigue: For example Reyner & Horne (1997) reported an increase of one step of the KSS after 10 minutes, while Moller, Kayumov & Shapiro (2003) observed microsleep events during a 30-minute test drive with non-sleep deprived subjects. In this regard, simulator experiments appear an economic way to examine continuously measured physiological parameters over a broad range of wakefulness steps (Moller et al., 2003). Like most subjective scales, these are rather ordinal in nature, but still more sensitive than performance indicators in low-workload tasks or when future performance changes need to be assessed (for a review see Annett, 2002). The latter is pivotal for the ultimate goal of a sleepiness warner that should issue a warning *before* performance breakdown occurs.

1.2 Oculomotoric parameters indicating drowsiness

1.2.1 Blink parameters. Blink frequency and duration as well as size of lid cleft (the distance between the upper and lower eyelid) remain the best-examined oculomotoric indicators for current alertness state and a driver's ability to react to environmental stimuli (Hargutt, 2003; Lal & Craig, 2001a; Papadelis et al., 2007; van Orden et al., 2000). It is well-known that fatigue is associated with increased blink frequency (Galley et al., in press; Hargutt, 2003; Hoffman, 1946; Luckiesh & Moss, 1937; Morris & Miller, 1996; Sirevaag & Stern, 2000; Stern et al., 1994; Summala et al., 1999), and blink duration (Caffier et al., 2003; Galley & Churan, 2002; Galley et al., in press; Häkkänen et al., 1999; Hargutt, 2003; Morris & Miller, 1996; Sirevaag & Stern, 2000; Svensson, 2004; van Orden et al., 2000; Verwey & Zaidel, 2000).

Hargutt (2003) investigated blink behaviour of eleven subjects in a driving simulator course, dividing sleepiness into four stages: awake, reduced vigilance, fatigued and sleepy. Light fatigue (the transition from awake to reduced vigilance) was indicated by an increase in blink frequency alone, whereas the transition to severe sleepiness was accompanied by an increase in blink duration. In an extensive literature review Meinold (2005) also concludes that blink frequency and duration may have to be considered as independent aspects of blink behaviour. An increased blink rate in connection with fatigue may be best understood as a cessation of attention-driven inhibition of blinks. Prolonged duration in contrast would reflect deactivation and slowing down of several physiological processes caused by decreased neuronal firing rates in the nervous system.

A lid closure that lasts more than about 500 milliseconds and covers the pupil for that time is usually defined as a micro*sleep*. However, a strict correspondence to traditional EEG-based sleep stages is still not fully substantiated – partly due to the fact that these short events do not correspond to the time-windows used for EEG analysis in sleep research. Harrison & Horne (1996) for example recommend time windows of at least 5 seconds for sleep diagnosis, but these are still too long to probe microsleeps events seen in sleepy drivers. For example, they averaged about 0.7 seconds in a study by Summala et al (1999) using video-analysis. Even the microsleeps of co-drivers that lasted up to 2.6 seconds in that study do not extend over the required time.

Given that long duration blinks usually occur repeatedly during severe sleepiness, the PERCLOS device measures the proportion the pupil is covered by the upper lid more than 80% in a 1 minute time-window (Dinges et al., 1998; Skipper & Wierwille, 1986; Wierwille, 1999) and yields a warning when a certain threshold is exceeded. Practical experience however suggests that not all sleepy drivers show overlong lid closures and some manage to keep their eyes open throughout the episode. According to O'Hanlon and Kelley (1977) some subjects even exhibit sleep spindles and K-complexes typical for sleep stage 2 in the EEG while driving and having their eyes open (quoted from Lal and Craig, 2001b). At this stage, adequate processing of environmental stimuli (e.g. reaction to suddenly appearing obstacles on the road) is almost impossible; however, while on an unrestricted road such drivers are presumably still able to keep their lane. To what extent unbraked rear-end collisions, especially at night, can be accounted for by this phenomenon is unknown. It is also described in literature as *driving without awareness* (Karrer et al., 2005), but little is known about its frequency, its relation to microsleeps and whether it occurs predominately in certain subgroups such as professional drivers. In any case, registration of lid closure alone is obviously insufficient to detect severe sleepiness in all people.

1.2.2 Saccades. Unlike lid cleft and blink duration, saccadic parameters have still not been considered as possible candidates for a sleepiness warner, although it is well known that *saccadic speed,* for example, is a reliable indicator of fatigue (Galley, 1989, 1993, 1998; Sirevaag & Stern, 2000). This measure has also been used in pharmalogical research to investigate fatigue-inducing effects of drugs or alcohol (Abel & Hertle, 1988; Blom et al., 1990).
The time between two saccades is generally called *fixation duration*. This event is closely related to cognitive processing in alert subjects (for a review see Findlay & Walker, 1999 , for a physiological modell see Trappenberg *et al.*, 2001, for application in vehicles see Victor *et al.*, 2005), but it has failed to show an unequivocal relation to sleepiness (Galley & Andrés, 1996; Lavine et al., 2002; Saito, 1992). This is surprising and may be due to the fact that fixations of different lengths may reflect different neuronal processes as seen in various studies (Godijn & Theeuwes, 2002; Pannasch et al., 2001; Radach et al., 1998; Unema et al., 2000; Velichkovsky et al., 2005): Indeed, very short fixations (< 150 ms), so-called express-fixations may turn out to be a distinct category caused by low level visuomotor behaviour; it could represent reflexive unconscious or non-cognitive aspects of behavioural control. We take up this suggestion and contrast the proportion of cognitive saccades (between 150 and 900 ms) with that of the very short (<150) the overlong (>900) saccades.

Registering saccades places higher demands on technical equipment than the recording of blinks: the required sampling rate of 500-1000 Hz cannot be achieved by most commercially available video systems today, but it is undisputed that any in-car warning device will need to be contactless. Therefore, neither the Electrooculogram (EOG) nor the most important biosignal in sleep research, the EEG (Pivik, 1991; Tassi & Muzet, 2001) will be a tenable option (Lal & Craig, 2002). Accordingly, the solution to this problem will most likely involve cameras. However, besides low sampling rates, tracking the eyes continuously under field conditions (i.e. in the car) remains a great challenge for such camera systems. In addition, overlong lid closures in sleepy subjects have to be distinguished from events like head turns or facial self-stimulation, where the eyes are not trackable by the sensor and image processing algorithm.

1.2.3 Interindividual differences. Interindividual differences in sleepiness behaviour were rarely taken into consideration in earlier studies (Van Dongen et al., 2004) and have only recently been declared a relevant field of research (Van Dongen et al., 2005). For reasons of small sample sizes alone, most results concerning interindividual differences (or better *homogenity*) in experimental sleep research have to be interpreted with caution. At the same time, individual deviations from group trends are of utmost relevance for a warning device, given that they may indicate points of failure in individual cases.

Summing up, the aim of the present experiment is to examine a variety of oculomotoric parameters in inviduals passing through a broad range of subjective alertness. This will be carried out in a large number of subjects in order to examine our parameters with regard to interindividual differences. Of special interest will be stages of light and medium fatigue that do not express themselves in clear performance detoriation at that moment, but which might still lead to a later performance breakdown that should be prevented by a sleepiness warner early enough. Thus the emphasis is primarily on subjective fatigue and its relation to numerous saccadic and blink parameters registered with high precision.

2. Methods

2.1 Participants and design

In two experiments a total of 138 subjects (76 males, 62 females, mean age 33,4 ± 11,5) were recruited and paid for their participation, resulting in 129 usable recordings. All participants were initially screened during a telephone interview for sleep disorders and were asked to refrain from consuming coffee or alcohol before the experiment. After instructing the participants and applying electrodes, the subjects were asked to drive a monotonous simulator course. The equipment was located in a dimmed laboratory/simulator room and they were asked to drive for about two hours (134 minutes on average), during which any road exits had to be avoided and their speed had to fall under a certain limit; leaving the road yielded a rattling sound. When the subjects reached an insurmountable level of sleepiness including repeated microsleeps or driving errors the test drive was stopped, otherwise it was continued up to 2 ¾ hours. Data were collected at two different locations and in two periods but were pooled because no meaningful differences existed between both data sets.

2.2 Data recording

The subject's face was filmed by a video camera and monitored by an experimenter sitting in an adjoning booth. Any objective indicator of fatigue like overlong lid closures (termed *microsleeps* from now on) or staring (termed *driving without awareness [DWA]s* from now on) was marked online via a PC-keyboard in the data stream. This online-coding was extended by an off-line video rating classifying the driver's face according to four stages of fatigue (see *data analysis*). The experimenter also coded online every obvious driving error as indicated by the rattling sound online.

The horizontal and vertical Electrooculogram (EOG) was recorded with an amplifier from PAR-Elektronik Berlin (0-250 Hz bandwith) and stored on the hard-disk of 486 DOS personal computer using a 12-Bit AD-converter with a sampling rate of 1000 Hz. Prior to recording, horizontal and vertical gazes were calibrated by letting the subject look at pre-defined spots in his/her field of vision in order to determine what changes in AD-values correspond to what angle of vision.

Commencing with the start of the test, the drivers were asked every thirty minutes to rate their subjective alertness on a scale from 10 *fully awake* to 1 *absolutely tired*. In addition, they were encouraged to report spontaneously perceived change in subjective alertness. A final rating was obtained immediately before stopping the driving course.

2.3 Data analysis

2.3.1 Blinks and saccades. The EOG raw data were analysed offline with a MATLAB®-based program that detects saccades and blinks automatically (Hofmann, 2005). In general, all saccades > 1° as well as all blinks are identified, the latter having a distinct shape in the vertical EOG. After smoothing the raw signal by zero-phase digital filtering, the underlying algorithm first determined the beginning and end of a saccade or lid closure/reopening by scanning the raw signal for step-like changes exceeding a minimum speed. This minimum speed was set to 30°/sec to separate them from slow signal drifts due to head movements or changing skin resistance. Next, possible saccadic or blink events were checked for plausibility according to several criteria like minimum duration, maximum speed etc[1]. Figure 1 shows a raw signal including identified saccades and blinks.

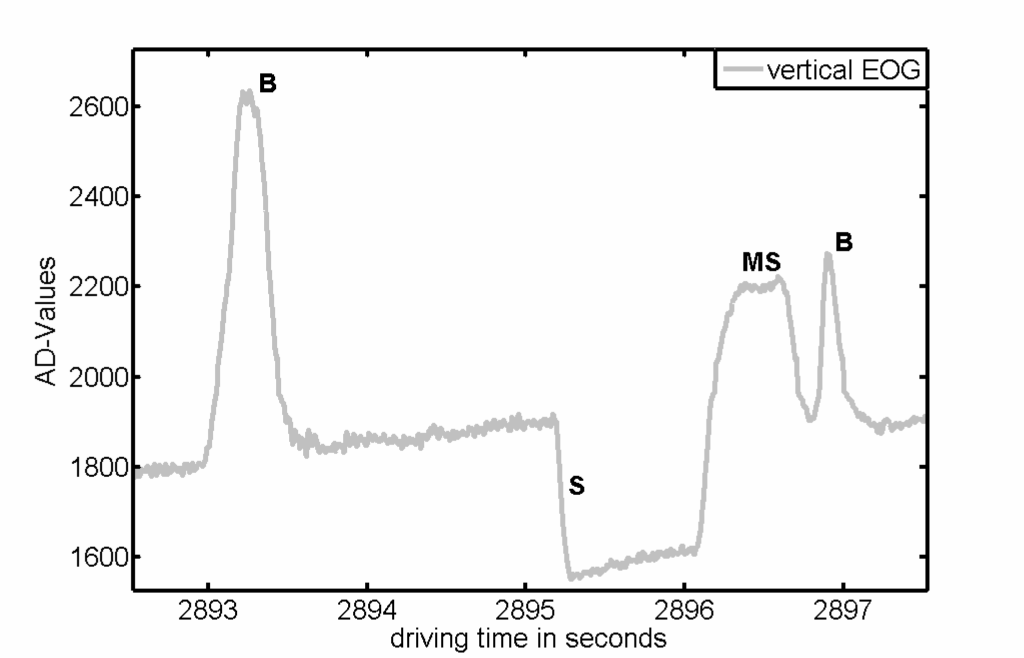


Figure 1: Vertical EOG (unfiltered raw signal) of subject 53 with two blinks (B), a saccade (S) and an overlong lid closure which meets the criteria of a microsleep (MS). In awake persons, lid closure (upward signal change) is almost immediately followed by its reopening (downward signal change) resulting in a spike-like artefact as in the second blink around second 2897. The first blink around second 2893 shows a slight delay between closing and reopening, giving the impression of a 'rounded' tip. During a mircosleep, this delay is substantially prolonged. Saccades start and end abruptly and are therefore easily recognized by detection algorithms. The microsleep includes an upwards saccade that cannot be separated from the lid closure movement.

For each valid event, the following parameters were determined:

- saccadic respective blink interval as the time from the end of the previous event to the beginning of the current event.
- saccadic duration as time from beginning to end of current event. Blink duration was defined as the time from start of lid closure to the moment of maximum speed during reopening, since the upper lid reaches its final position asymptotically in the reopening phase and precise stopping is hard to determine unequivocally; the moment of maximum speed is always unambiguous.
- saccadic respective lid closure amplitude
- saccadic respective lid closure speed, expressed as average as well as maximum speed
- delay of reopening, defined as the time after full closure to the start of reopening the upper lid eye lid. In awake subjects, this delay last only a few milliseconds but is increased under fatigue and reaches values of several hundred milliseconds during a microsleep.

For both saccades and blinks, duration and speed depend on the respective amplitude (Galley & Andrés, 1996); therefore *standardised* duration and speed were also computed and expressed as percent of expected duration or speed for the respective amplitude. These standardised parameters were corrected for possible covarying changes in amplitude and thus represent 'pure' central nervous activation. The

---

[1] As blinks were defined as an upward saccade immediately followed by a downward saccade, overlong blinks with a delay of more than 100ms between full closure and reopening are difficult to distinguish from an upward gaze followed by a downward gaze. For that reason, some microsleeps events marked by the experimenter were missed by the algorithm.

underlying formula were derived using data from our laboratory following theoretical assumptions by Becker (1989) and Collewijn *et al.* (1988). For a different approach see Hargutt (2003).

2.3.2 Video rating. To obtain a second sleepiness measure independent of subjective self-ratings, an extensive drowsiness evaluation based on video-recordings of the driver's face (Vöhringer-Kuhnt et al., 2004) was performed for 64 subjects of the second data set. The videos were divided into segments of one minute and each segment was assigned to one of the following fatigue stages: 1 = not drowsy (fast blinks and saccades, normal facial tonus), 2 = drowsy (frequent blinks, limp face, self-activating behaviour like yawning, scratching), 3 = very drowsy (clearly prolonged lid closures, rare blinks, staring or drifting eyes), 4 = extremely drowsy (overlong lid closures > 2 seconds, microsleep attacks and scaring up abruptly). The interrater-reliability of this procedure is r=.712 (Vöhringer-Kuhnt et al., 2004). This off-line video rating of facial behaviour was then compared to self-rated alertness.

2.3.3 Subjective alertness ratings. The alertness self-ratings taken every 30 min or reported spontaneously were interpolated linearly with respect to time so that each blink and saccade could be labelled with a specific subjective alertness value. This was based on the assumption that wakefulness decreases with increasing time-on-task and time-since-sleep (Sirevaag & Stern, 2000). For all oculomotoric events of the 64 subjects mentioned above there was also an interpolated video-rated alertness value.

2.4 Data aggregation

Finally, all data for each subject were aggregated over a time window of five minutes, computing the mean of self-rated as well as video-rated alertness and the mean, median, and standard deviation for each oculomotoric variable. If a standardised version of an oculomotoric variable was available, there were in total 2*3 statistical parameters depicting the initial oculomotoric measure. For example, blink duration could be expanded to yield six parameters: mean(duration), median(duration), SD(duration), mean(standardised duration), median(standardised duration), SD(standardised duration).

Besides this time-based approach, all variables were aggregated over the subjective alertness steps for each subject; This provided data that were less susceptible to time-on-task effects. All data and statistical analyses were done with the SPSS® 12 and MATLAB® 7 software packages.

3. Results

3.1 Alertness ratings

Overall -alertness decreased substantially in subjects driving the two-hour simulator course. Subjective alertness ratings as well as video-rated drowsiness both showed an almost linear increase in sleepiness (see figure 2).

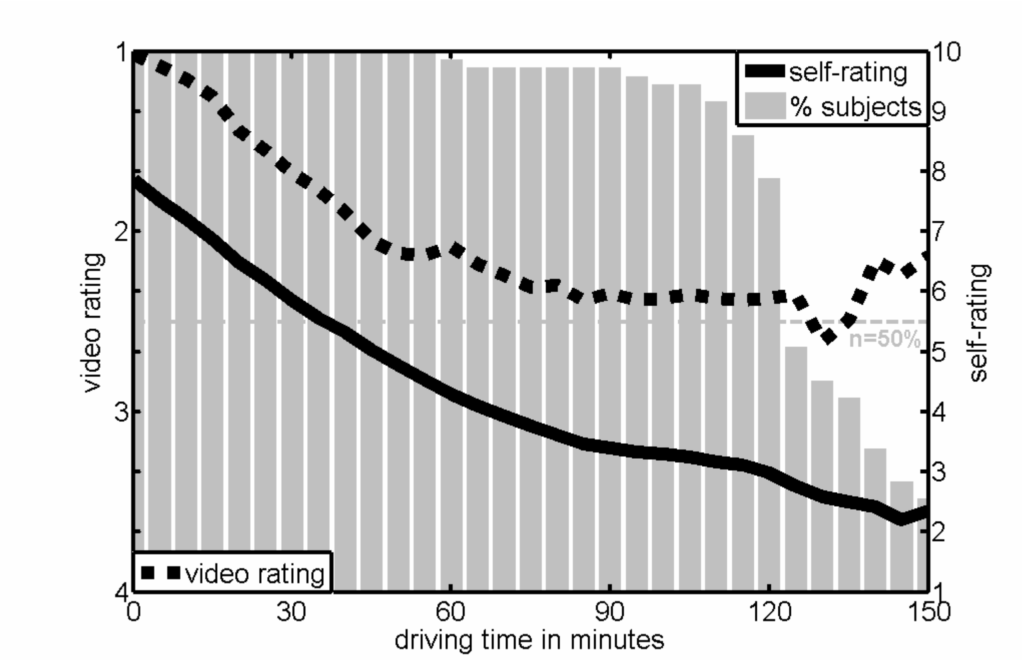


Figure 2. Average self-rated alertness and video-rated fatigue during simulator course. Grey bars depict proportion of subjects that reflected the respective mean rating: Full length means n=100%. Note that increased fatigue is scored by increasing values in the video rating whereas in self-rating by decreasing (alertness) values.

Average self-rated alertness dropped from 8 in the beginning to 3 after two hours driving time for the whole sample of n=129. When reaching critical values less than 3 the overall trend gets weaker as the experiment was stopped for extremely tired subjects that could not carry on driving. 84 subjects showed microsleeps, on average 26 of these overlong lid-closures (SD=30, median=13) per person. Video-rated drowsiness reached saturation after about only 90 minutes with an average value of 2.4 which corresponds to 40 percent of remaining alertness when interpolating the video fatigue steps linearly. To assess the correspondence of self-rated subjective and video-rated objective sleepiness, the 5-minute-means of the two variables were rank-correlated over all subjects, resulting in r=-.639 (1555 values of n=64, $p<0.001$).

The distribution of observed microsleeps over subjective alertness also shows a close relationship between the two measures of drowsiness (see figure 3): 87% of all microsleeps occurred at low stages of self-rated alertness, 48% at the lowest stage. Microsleeps were rarely seen when subjects rated themselves as being alert.

The same holds for driving errors: 51% occurred at alertness level 1. Altogether 84% of all errors were made at the stages of high subjective sleepiness (alertness level 1-3).

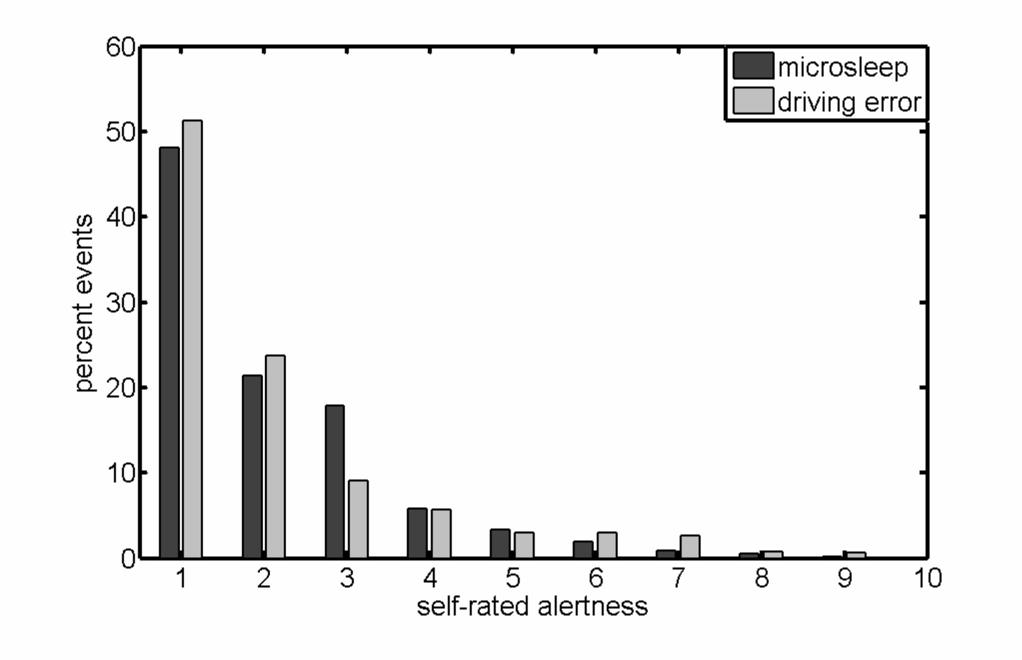


Figure 3. Microsleeps and driving errors abruptly increase at stages of low self-rated alertness. Note that increasing alertness from left to right on the x axis reverses the actual temporal course of the experiment where subjects started with high alertness values and alertness dropped over time.

3.2 Oculomotoric parameters as indicators of subjective and objective sleepiness

After data aggregation several correlations were determined to identify oculomotoric variables that are sensitive to fatigue-related changes. In sum, there are two time-based approaches that use data aggregated over 5-minute-windows or evaluate changes with respect to increasing fatigue over time. A different approach is exclusively based on subjective fatigue-steps with no relation to driving time. These correlations were derived as follows:

- For each subject (n=129) an individual correlation was determined between average subjective alertness during a 5-min time-window and the statistical parameters (mean, median and standard deviation) of each oculomotoric variable in this time-window. On average, this individual correlation was based on 25 time-window values. Individual correlations were Z-transformed, the Z-values averaged and re-transformed into a global correlation that represents the mean individual correlation between the mean, median or standard deviation of an oculomotoric variable and subjective alertness. The degrees of freedom for the significance level were set to df=25 based on the aforementioned 25 5-min-time-windows for each person.
- For the subjects that additionally had been video-rated (n=64) an individual correlation was determined between average video-rated alertness during a 5-min time-window and the statistical parameters (mean, median and standard deviation) of each oculomotoric variable in this time-window. On average, this individual correlation was also based on 25 time-window values. Individual correlations were Z-transformed, the Z-values averaged and re-transformed into a global correlation that represents the mean individual correlation between the mean, median or standard deviation of an oculomotoric variable and video-rated alertness. The degrees of freedom for the significance level were set to df=25 based on the aforementioned 25 5-min-time-windows for each person.
- One global correlation across all subjects (n=129) was determined between non-interpolated subjective alertness steps and respective aggregated oculomotoric variables, in order to check for a possible bias caused by the interpolation. As each subject passed through about 6 alertness steps, an individual approach was not feasible here because correlations of single subjects were -1 or +1 in many cases.

Thus, global correlations of altogether 780 individual alertness ratings (≈ 6*129) with their oculomotoric counterparts were used.

For each approach, the correlations were ranked according to their absolute value.

Table 1. Spearman Rank Correlations between oculomotoric parameters and subjective respective video rated alertness.

Oculomotoric variable (*statistical parameter*) sd: standard deviation stand: standardised max: maximum	Mean individual correlation with subjective alertness in 5min-time-windows (129 subjects)		Mean individual correlation with video-rated alertness in 5min-time-windows (64 subjects)		Global Correlation with subjective alertness steps (129 subjects)	
	rho	Rank	rho	Rank	rho	Rank
Blink duration (mean)	**-.582****	1	**.684****	1	**-.331****	5
Standardised blink duration (mean)	**-.557****	2	**.564****	3	**-.341****	4
Blink duration (sd)	**-.519****	3	**.610****	2	**-.390****	2
Standardised blink duration (median)	**-.506****	4	**.525****	5	**-.302****	6
Blink duration (median)	**-.504***	5	**.540****	4	**-.291****	7
Standardised blink duration (sd)	**-.480***	6	**.481***	9	**-.418****	1
Delay of lid reopening (mean)	**-.478***	7	**.522****	6	**-.275****	8
Blink interval (median)	**.453***	8	**-.408***	12	**.148****	27
Standardised lid closure speed (mn)	**.449***	9	**-.494***	7	**.275****	9
Delay of lid reopening (sd)	**-.429***	10	**.486***	8	**-.352****	3
Standardised lid closure speed (median)	**.426***	11	**-.464***	10	**.272****	10
Saccadic duration (sd)	**-.383**	12	**.440***	11	**-.263****	11
Blink interval (mean)	**.379**	13	**-.303**	23	**.061**	57
Delay of lid reopening (median)	**-.350**	14	**.390**	14	**-.208****	14
Saccadic duration (mean)	**-.347**	15	**.336**	17	**-.126****	38
Lid closure speed (median)	**.335**	16	**-.393**	13	**.199****	15
Lid closure speed (mean)	**.330**	17	**-.383**	15	**.186****	20
Standardised saccadic duration (mean)	**.313**	18	**-.360**	16	**.092****	50
Standardised saccadic duration (median)	**-.272**	19	**.312**	21	**-.082***	55
Standardised lid reopening duration (mean)	**-.268**	20	**.315**	19	**-.198****	16
Saccadic amplitude (sd)	**-.268**	21	**.226**	34	**-.144****	30
Stand. lid reopening duration (median)	**-.260**	22	**.326**	18	**-.218****	12
Standardised saccadic duration (mean)	**-.257**	23	**.314**	20	**-.123****	38
Standardised lid closure speed (sd)	**-.240**	24	**.212**	38	**-.181****	21
Standardised max. lid closure speed (mean)	**.237**	25	**-.291**	25	**.160****	25

Subjective alertness was asked every 30 minutes and interpolated for correlations in columns 2 and 4 (see section 2.3.3). Negative correlations in column 2 and 6 mean that values of oculomotoric variables (e.g. blink duration) increase with decreasing subjective alertness. Subjective and video ratings utilized opposite encodings of fatigue, therefore the correlations in column 4 show the opposite signs as those in column 2 and 6.

First it is noteworthy that both measures of fatigue suggested the same oculomotoric variables as possible indicators of fatigue. Although based on different criteria (self-rating or video-rating), the ten highest rankings are almost identical in both correlation lists.

Blink duration was by far the most important variable to indicate both subjective and objective sleepiness. Standardised with respect to blink amplitude or in its unstandardised version, the statistical parameters mean, standard deviation (sd) and median of blink duration occupied the first five places in both rankings. Next the *delay of lid reopening*, followed by *lid closure speed* and *blink interval* were most sensitive to changes in fatigue.

Looking at saccadic parameters, *saccadic duration* as well as *saccadic speed* and *saccadic amplitude* also showed a relationship to sleepiness that was statistically significant for global assessment (one global correlation across all subjects, table 1 column 6).

In general, the correlations of oculomotoric parameters with video-rated alertness were slightly higher than with subjective alertness; this can be explained by the fact that the latter was asked every thirty minutes whereas the video was rated continuously. At the same time, averaged individual correlations (columns 2 and 4 in table 1) lead to remarkably higher values than global correlations across all subjects (column 6 in table 1). This provided the first indication of substantial individual variations. It needs to be noted that the global correlations which were based on non-interpolated alertness ratings also showed the same ranking of parameters for the first ten variables (with the exception of blink interval which is extremely affected by inter-individual variation); therefore, the interpolation apparently did not lead to a misrepresentation of underlying relations.

Unexpected to us, increases in individual standard deviations during 5-minute-windows are about as highly correlated with sleepiness as increases in mean, in some cases even higher (e.g. saccadic duration). Thus, for some variables a rise in variance seems to be a better indicator of fatigue than changes in mean or median. Qualitatively, increased fluctuations represent something very different than just decreasing vigilance. We will interpret this as intensified attempts to counter-regulate with growing sleepiness (see *discussion*).

3.2.2 Parameter course over subjective fatigue. In order to check whether changes in oculomotoric variables due to fatigue are continuous and linear, their statistical parameters were plotted against subjective fatigue steps. Again separate means, medians or standard deviations for each subject were initially calculated; These individual values were then used to determine the global statistical parameter. This stepwise procedure was chosen to ensure that each subject contributed equally to the overall value regardless of how many blinks/saccades she or he made. Figure 4 shows the increase in mean and median blink duration, the best indicator of sleepiness with decreasing alertness.

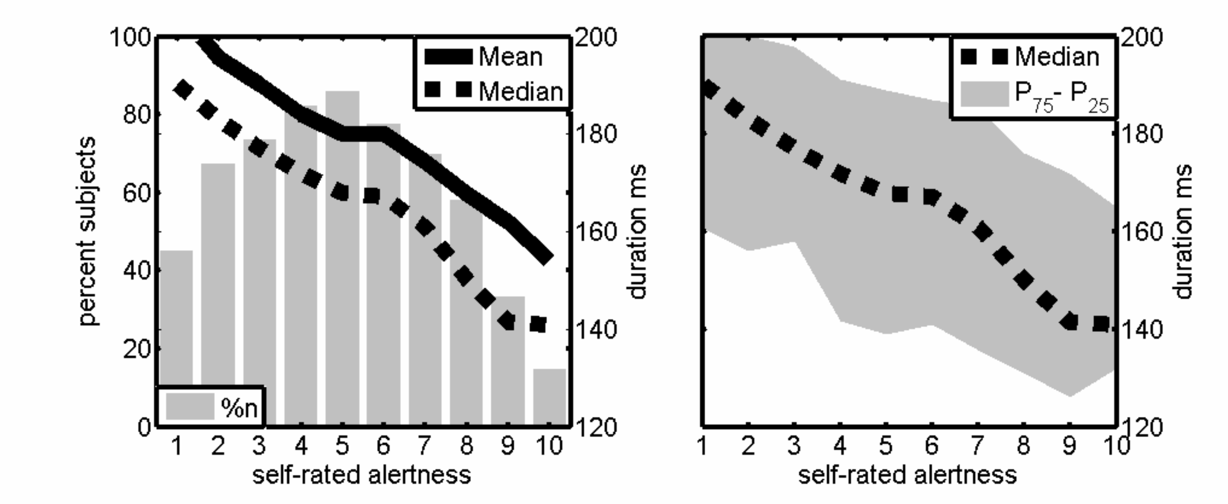


Figure 4. Blink duration over self-rated alertness. Left: mean and median (solid and dotted line respectively). Grey bars indicate percentage of subjects that contributed to each alertness step. Right: overall median (dotted line) and the 50% range of individual medians (grey band) per alertness step.

Both parameters revealed a continuous rise with decreasing alertness. The additionally depicted percentage of subjects per fatigue step (grey bars figure 4 left) indicated sufficient sample sizes for all stages of light and severe fatigue even including the lowest alertness level (note the aforementioned drop-out effect for extremely drowsy drivers); However, there was an underrepresentation of stages of high alertness – only a few subjects rated themselves as absolutely awake at the beginning of the experiment. From the right part of figure 4, one notes the substantial deviation of individual values from an overall median: the grey band illustrates the 50%-percentile range of individually derived medians for each alertness step. This reveals that even during alertness level 9 individual medians for 50% of the whole sample already vary from 130-170 ms whereas the overall median is 140 ms.

From table 1, delay of reopening turned out as the second-best indicator of fatigue. Figure 5 shows the gradual increase of this parameter whose values are in the range of millicseconds. Although the median is insensitive to increases that occur during light fatigue, it suddenly changed when reaching alertness stages <4, a level that was associated with severe fatigue i.e. cumulative occurrence of microsleeps. For this variable, individual variation apparently increases with decreasing alertness as the grey band depicting the 50%-range of individual means (figure 5, right) broadens with lower alertness ratings.

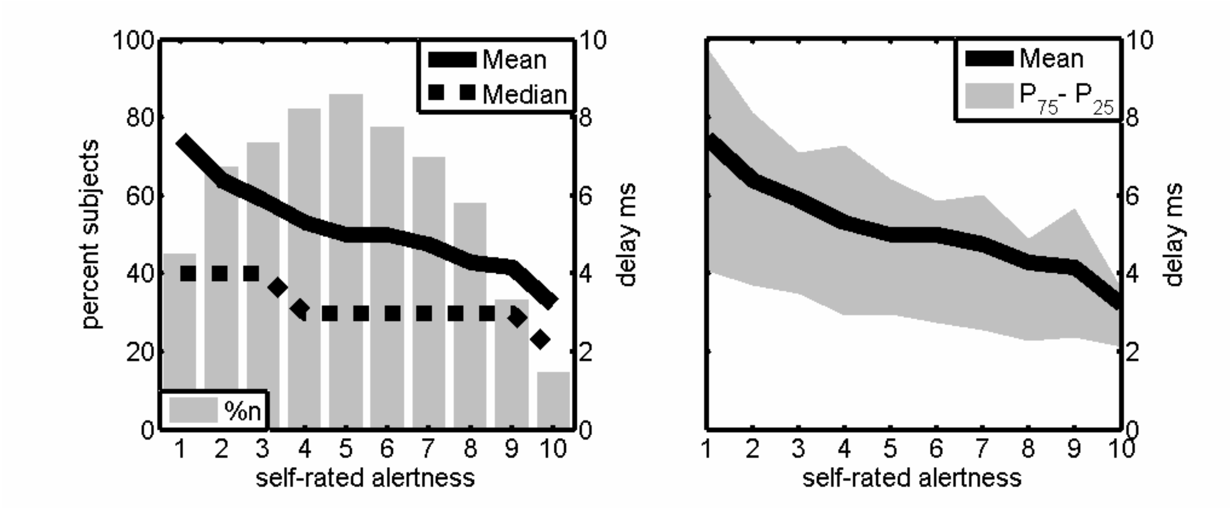


Figure 5. Delay of reopening over self-rated alertness. Left: mean and median (solid and dotted line respectively). Grey bars indicate percentage of subjects that contributed to each alertness step. Right: Overall mean (solid line) and the 50% range of individual means (grey band) per alertness step.

As expected, blink interval decreased with growing fatigue. However, the inconsistent course of mean interval compared to median interval showed that this variable was susceptible to outliers and strong variation which affect the median less than the mean. Nevertheless, individual variations outnumbered general decreases due to fatigue (see figure 6, right).

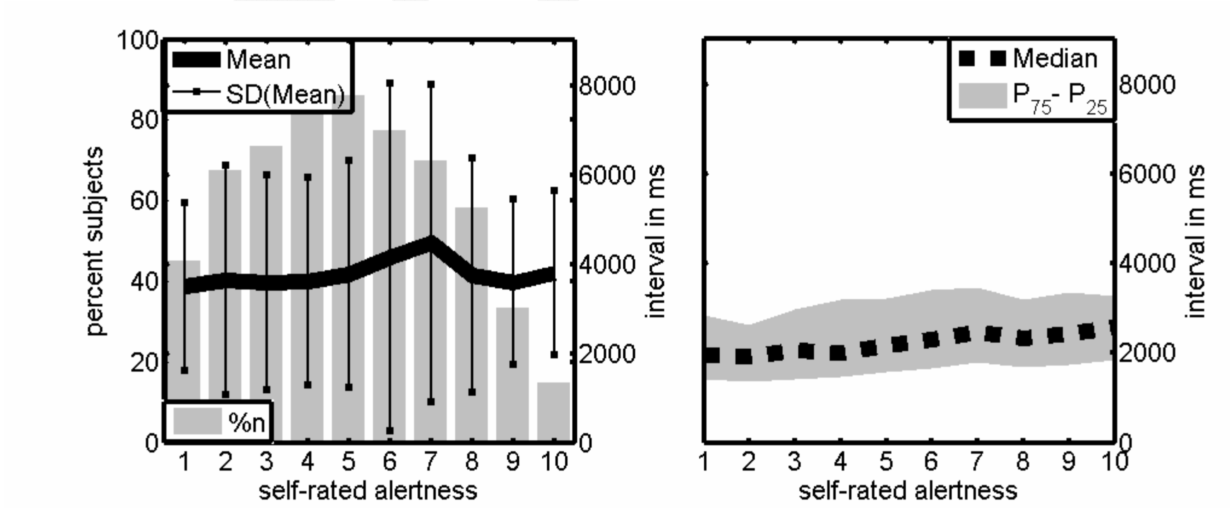


Figure 6. Blink interval over self-rated alertness. Left: Overall mean and variation of individual means (bold and thin black lines respectively). Grey bars indicate percentage of subjects that contributed to each alertness step. Right: Overall median (dotted line) and the 50% range of individual medians (grey band) per alertness step.

Next to these 'classical' oculomotoric indicators of sleepiness we also addressed two less considered variables. These were *standardised lid closure speed* and *saccadic duration*.

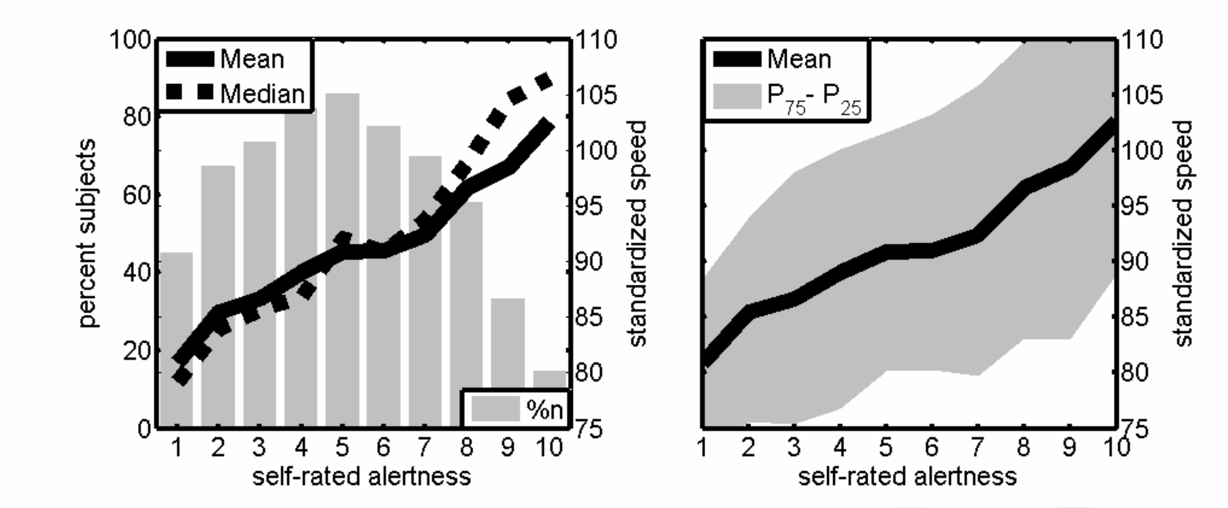


Figure 7. Standardised lid closure speed over self-rated alertness. Left: mean and median (solid and dotted line respectively). Grey bars indicate percentage of subjects that contributed to each alertness step. Right: Overall mean (dotted line) and the 50% range of individual means (grey band) per alertness step.

As described in the methods section, lid closure speed was standardised with respect to amplitude to eliminate any possible impact of amplitude changes. After this adjustment, standardised lid closure speed shows a clear deceleration with increasing sleepiness (see figure 7) while this effect is less obvious in unstandardised closure speed (rank 16-17 in table 1).

Besides characteristics of lid movement, we also considered saccadic parameters as possible variables for drowsiness detection. Of all saccadic variables, standard deviation of saccadic duration showed the highest correlation to video-rated fatigue (r=.440, see table 1). Again, standard deviation was determined on an individual level to account for interindividual differences.

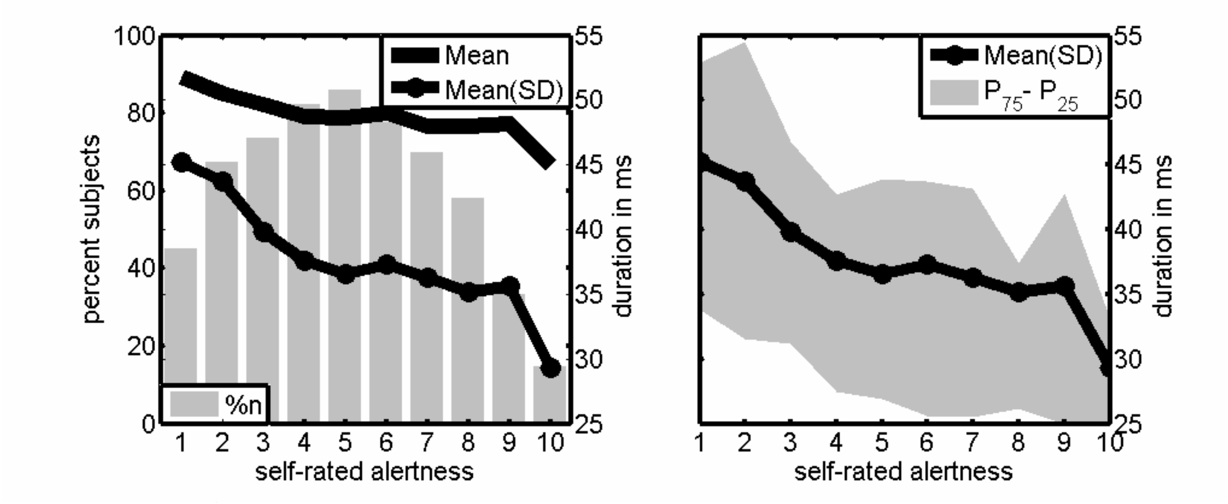


Figure 8. Saccadic duration over self-rated alertness. Left: mean saccadic duration (solid line) and averaged individual standard deviation (mean(SD), solid line with black circles). Grey bars indicate percentage of subjects that contributed to each alertness step. Right: Averaged individual standard deviation (mean(SD), solid line with black circles) and the 50% range of individual standard deviations (grey band) per alertness step.

Figure 8 depicts the global mean saccadic duration that increases only moderately with decreasing subjective alertness. Averaged individual standard deviation however increases notably, especially during stages of low alertness (ratings < 4). During severe sleepiness, drivers apparently stop monitoring their environment in a constant way and start to look around unsystematically; this results in saccades with greater variability in amplitude and duration up to the oculomotoric correlative of 'scaring up' after lapses of attention, i.e. looking around to re-orientate oneself.

3.2.3 Fixation duration. Consistent with literature, mean fixation duration showed *no relation* to fatigue. Taking a closer look, we divided fixation duration in three classes, short (<150 ms), middle (150-900 ms) and overlong (>900 ms) fixations and examined the course of these different classes representing different mental processes during decreasing alertness (see figure 9).

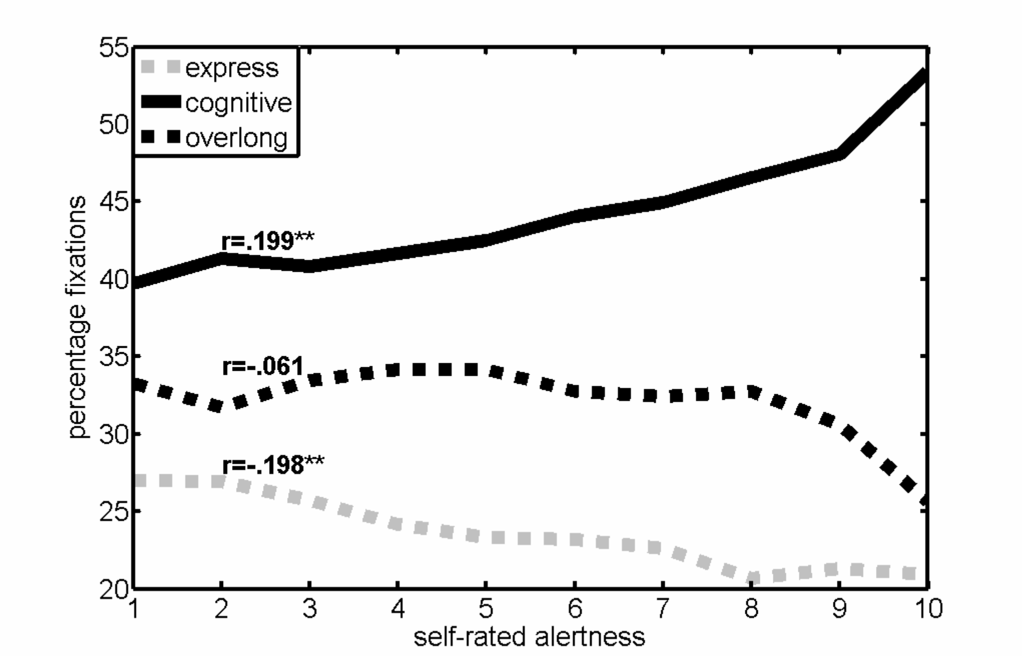


Figure 9: proportion of different types of fixation durations during decreasing subjective alertness steps. During wakefulness cognitive fixations with a duration between 150 and 900 ms dominate, while with increasing sleepiness express (<150 ms) and overlong (> 900 ms) fixations become more frequent, accompanied by a relative decrease of cognitive fixations. r = rank correlation between percentage of fixation type and alertness rating; **= p<0.01 (780 values of n=129).

The different types of fixations show distinct changes due to fatigue: the proportion of fixations with a duration between 150 and 900 ms that is associated with cognitive processing clearly decreases, while the percentage of express and overlong fixations (staring) increases. On the level of average fixation duration, though, these contrary courses tend to neutralize each other. Albeit the correlation between proportion of cognitive fixations and alertness is only moderate (r=.199, p<0.001, 780 values of n=129), it can be assumed that with growing sleepiness the subject shows decreasing interest in the environment (see *discussion* and *appendix B*).

3.2.4 Blinks around a microsleep. As the observed microsleeps had an unique timestamp, we were able to analyse saccades and blinks immediately before and after an overlong lid closure in 80 of the 84 drivers that showed at least one microsleep (each of these subject had 26 microsleeps on average). For an example of blinks before and after a microsleep see figure 1. Figure 10 illustrates that microsleep-blinks with a median length of 269ms were substantially longer than the blinks immediately before (median 204 ms) and after (median 189 ms). At the same time, the median value of 269 ms is lower than the expected 500 ms noted in the literature (see introduction). This suggests that the video-rating clearly detected longer blinks but had already assessed shorter ones as microsleeps. Moreover, the blink-detection algorithm might have misinterpreted some of the very long lid closures as two separate saccades (see figure 1).

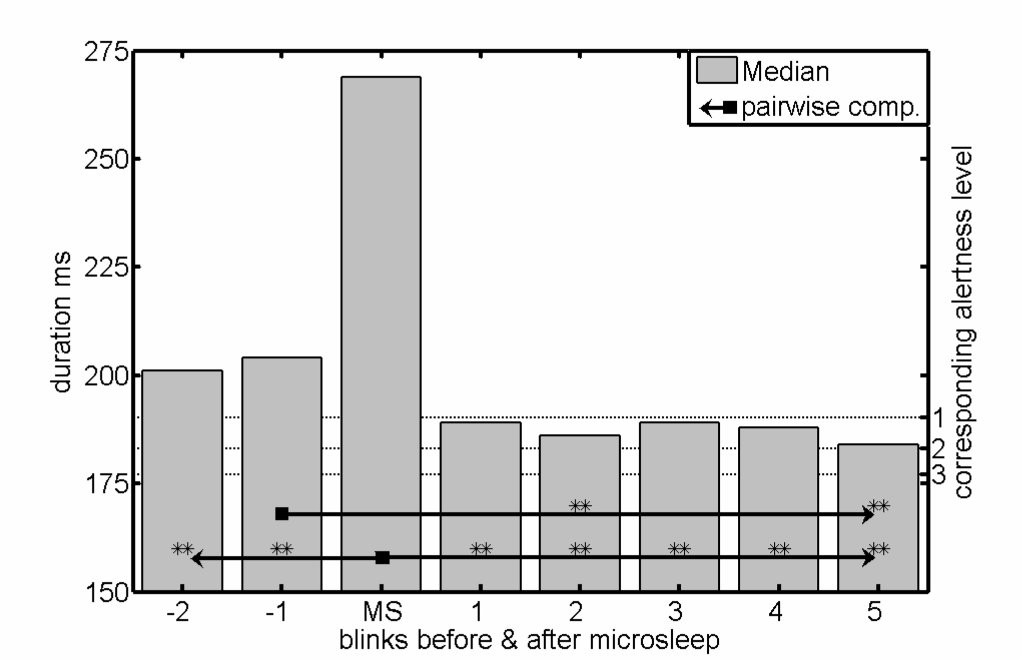


Figure 10: Blink duration before and after an overlong lid closure coded as a microsleep.**= $p<0.01$ in pairwise comparism.

The dotted lines in figure 10 also show that subjective alertness remains very low before and after a microsleep. Blink duration is only slightly albeit significantly decreased after a microsleep; this would actually argue that there is a little less fatigue after a microsleep (see discussion).

4. Discussion

4.1 Subjective and objective sleepiness

The monotonous simulator course successfully induced severe objective and subjective sleepiness (see section 2.2) in most subjects (figure 2). In 84 of 129 subjects overlong lid closures, so-called microsleeps, were observed. They were also apparent in the off-line video rating of facial behaviour, which was proposed as an objective sleepiness score by Wierwille and Elsworth (1994). Interestingly, 87% of these microsleep events occurred during subjective alertness level 1-3 and 48% exclusively during alertness = 1. This was quite similar to the pattern seen for driving errors and validates self-rated alertness as a reliable measurement. These findings clearly contradict the common disregard for subjective sleepiness in many studies that fail to collecting it. Keeping in mind that subjective and observable sleepiness represent different aspects of fatigue (Ahsberg et al., 2000) and were originally collected with quite different time scales (frame by frame video-analysis vs. 30 minutes intervals), the correlation of .639 (see 3.1) between both ratings has to be considered as fairly high. Consistent with this positive assessment of subjective scores of alertness, the comparable nine-step *Karolinska Sleepiness Scale (KSS)* has repeatedly yielded meaningful results (Horne & Baulk 2004, Inger et al 2006 a,b). In our study only two subjects showed an obvious discrepancy between self-rated and video-rated alertness. Of these two, one subject rated his alertness >= 4 despite 21 microsleeps, but this person started with a value of 10, indicating a decrease of 6 alertness steps.

In the beginning, subjective alertness dropped on average one alertness step every 20 minutes (see figure 2). A comparable, albeit stronger increase of sleepiness was reported by Reyner and Horne (1997) with one step of the KSS every 10 minutes of driving the simulator. Both rates are actually too high to be solely explained by variables such as preceding time-since-sleep or previous strain in the context of circadian fatigue as attempted by existing sleep-wake-models (Mallis et al., 2004). Indeed, our clear time-on-task effect needs further explanation and can only partially be accounted for by situational monotony. It is a well-known fact that whereas sleepiness arises slowly during real-life rides, it arises more quickly under safe and monotonous laboratory conditions. In our view, this provides another example of an *appetence-function for sleep* (Lavie, 1991, 2001) that sometimes superimposes circadian rhythmicity and previous exhaustion.

From an appetence-model of sleepiness one could argue that a comfortable state of vita minima would be sought as soon as there is no other desirable goal to pursue or no possibility of altering the current situation like on long-distance motorway rides (Adams-Guppy & Guppy, 2003). According to Pinel (1997), satiated large mammals of prey are able to sleep up to 20 hours a day in the absence of natural enemies. Reitter (2001) reported that a monotonous car ride induces drowsiness faster if an unnoticed, but previously rated aversive high-frequency engine noise is present. In this context the author legitimately points to the fact that leisure activities like collecting stamps are highly fatigue-inducing for many people but not for the philatelist.

4.2 Oculomotoric indicators of sleepiness

4.2.1 Videorating and subjective sleepiness as independent variables. The correlations of oculomotoric variables with different indicators of sleepiness were comparable across all three correlation-approaches (average self-rated alertness in 5-min-windows, average video-rated alertness in 5-min-windows, aggregated over self-rated alertness steps without interpolating, see table 1). Rank correlations with video-ratings lead to slightly higher values than with interpolated self-rated alertness, which was most likely due to the fact that videos were rated continuously whereas subjective sleepiness was interpolated based on values asked every thirty minutes. Thus an interval of 15-20 minutes for collecting subjective ratings might be more appropriate than the 30 minute intervals used here. Aggregating on a trial basis yielded increasing overall-correlations for this criterion when extending the time windows from 1 up to 20 minutes. Thus, sleepiness is apparently a *tonic process* where temporary reawakening is possible. Raw signal (see figure 1) as well as averaged values (see figure 10) show that blinks after a microsleep are of shorter duration than the ones before, suggesting that a microsleep may induce a slightly higher vigilance level.

4.2.2 Blink parameters. Blink duration is by far the most important variable to indicate subjective as well as video-rated fatigue: their standardised and unstandardised version and their different statistical parameters occupy the highest positions in all correlation lists. Next come *delay of lid reopening*, *blink interval*, and *standardised lid closure speed*. Delay of lid reopening has only rarely been investigated as an indicator of sleepiness in literature so far (Caffier et al., 2003; Johns et al., 2007; Tucker & Johns, 2005)[2]. This is likely because of its need for a high sample rate: during wakefulness there is practically no delay between closing and reopening the eyelid and thus its values are in the range of 1-4 milliseconds (see figure 5) while during a microsleep delays of several 100 ms are possible. When completely falling asleep reopening is totally suspended, suggesting a close relationship between increasing delay of reopening and the urge to fall asleep. Accordingly, the reopening process might be coupled to a closing process by a discrete neuronal process, maintained only during wakefulness and then stopped during sleep. This parameter might represent a marker for the beginning of sleepiness and its end, sleep.

Lid closure speed decreases with increasing sleepiness; however, this has not yet been propagated as a possible indicator, which again may be due to insufficient sampling rate. For saccadic speed, there appears to be a direct correspondence to firing rates of oculomotoric neurons (Smit & Van-Gisbergen, 1990). Assuming the same for lid movements (which is not known yet, but see proposed similarities in Ongerboer de Visser and Bour 2006), then speed and duration would represent a direct link to central nervous activation. As both variables systematically vary with amplitude, standardised values should be preferred if one is interested in underlying neuronal activation (App & Debus, 1998; Galley, 1989). In

---

[2] Caffier *et al.* (2003) describe the parameter, but only report results for *blink duration* and *reopening time*

unstandardised versions, changes in amplitude can amplify or attenuate fatigue-related changes, depending on the particular direction of change. Blink duration is the sum of lid closure duration, delay of reopening and reopening duration (in our definition the time until the moment of maximum velocity during reopening, see *data analysis*) and thus made up by three parts that all increase with sleepiness. This is most likely one reason for its prevalence among other oculomotoric indicators of fatigue. Future drowsiness-detection algorithms, however, could utilize a decomposition in different sections in case certain phases of a blink event can be determined more reliable than others: start and end of closure phase can usually be specified more precisely than start and end of the re-opening phase.

Blink interval correlated unexpectedly low with alertness in our data set. Correlations with subjective alertness steps only reached rank eight, 13 and 27 for the median, mean and overall calculations, respectively (see table 1). There may be several explanations for this: besides considerable interindividual differences in blinking rate (the reciprocal of blink interval), a lacking sensitivity of blink rate for drowsiness in general may account for this result (Johns et al., 2007). Complying with the general view (for a review see Meinold, 2005) that increasing blink rate actually reflects a reduction of blink-inhibition caused by decreasing attentional demands or visual load, the monotonous experimental setup may have its effect well before sleepiness set in. In subjects that already show a high blinking rate, drowsiness can cause only minor increases. Moreover, several subjects began to stare during severe sleepiness and showed almost no blinks or saccades any more after an initial increase in blinks. This resulted in an inverted-U-shape relationship between blink frequency and sleepiness. Taken together, these opposed individual trends would lead to a low overall correlation.

4.2.3 Fixation duration. As reported, mean fixation duration appeared to show no relation to fatigue (Galley & Andrés, 1996; Lavine et al., 2002; Saito, 1992). However, after dividing fixation durations into different fractions of very short (< 150 ms), middle (150-900 ms), and overlong (>900 ms) fixations, clear changes due to fatigue became apparent: fixations of medium length showed decreases and at the same time the proportion of very short and overlong gazes exhibited increases. As the virtual scenery remained constant throughout the whole simulator course, changes in the visual environment as known from other studies can be excluded. Galley & Andres (1996) for example report that driving in city traffic resulted in average fixations of 350 ms while on the motorway the mean fixation was 450 ms. Here, the richer visual environment of a city causes shorter fixations, whereas the comparatively constant optic flow on a motorway mainly induces smooth pursuit eye movements and saccades are only necessary to fixate a new object that is passing by the visual field. Visual processing of complex scenes with rapidly changing stimuli (e.g. a city ride) leads to fixations typically between 200-400 ms, which is once again exceeded by fixation durations of around 250 ms during reading (Rayner, 1998). Fixations in this range are thus associated with content-related identification or *cognitive processing* and this type (ranged broadly from 150 ms to 900 ms in our study) represent about 50% of all fixations during wakefulness and still 40% during severe sleepiness. Shorter fixations (<150 ms) in contrast are insufficient to extract relevant information. In our introduction we quoted references that suggest a low-level processing during these express fixations, such as re-adjusting the gaze on the target after large saccades or sweep return during reading (Hofmeister et al., 1998). For overlong fixations (>900 ms) no common functional interpretation has been established yet, but in our study they were predominately related to stages of low alertness, i.e. staring during DWA or microsleeps. The decline of medium fixation durations that represent cognitive or attentional monitoring is accompanied by an increase of reflexive (very short fixations) or absent (overlong fixations) scanning behaviour with growing sleepiness. Given that this trend persists throughout all stages of decreasing alertness, we draw the following conclusion: increasing fatigue is continuously associated with decreasing interest in one's environment, which otherwise inhibits reflexive saccades (<150 ms) or disruptive spontaneous blinks and thus guarantees extensive visual processing during wakefulness. This *attentionial disinhibition* of reflexive saccades and spontaneous blinking apparently is an independent property of sleepiness that is reflected in increased blink rate and less cognitive saccades. *Deactivation* on the other hand becomes manifest in decreased oculomotoric performance, i.e. increased saccadic and blink durations and decreased speed. Preliminary factor analyses of our data also group saccadic and blink intervals on a distinct factor.

4.2.4 Standard deviations as an independent indicator of fatigue. For many variables, increases in individual variance correlated equally high or higher with fatigue than changes in mean or median. This demands further explanation, especially when variance increases whereas mean or median decrease as in speed-related-variables. Neither deactivation nor attentional disinhibition can account for increases in variability. At the same time it is obvious that while driving, sleepiness would be associated with a drive to stay awake. Self-activating behaviour was named as an indicator for stage 2 of drowsiness in the video-rating scale following Wierwille and Elsworth (1994). Subjective measures like the Karolinska Sleeping Scale also contain descriptions of effort or attempts to fight sleepiness, a phenomenon well-known in literature:

> "When habituation or decreased cortical alertness are combined with the opposing requirement to maintain a constant high level of alertness this results in considerable effort ... this may very well represent a combination that is quite stressful“ (Thackray, 1981).

Then again we found a slight increase in vigilance after a microsleep, depicted by slightly shorter blink durations. Taken together, increased variability eventually reflects a more or less successful attempt to fight sleepiness by the means of self-activtation. Higher fluctuation during sleepiness is mainly seen in oculomotoric variables representing general activation, namely durations and velocities. Less fluctuation is seen in attention-related factors like blink or saccadic interval. Next to attentional disinhibition and deactivation, increased attempts at this self-activation apparently comprises a third component of driver sleepiness which can also be observed in cardiovascular parameters (Kaida et al., 2007). Whether these attempts occur predominately in response to perceived performance errors or follow ultradian circles has to be examined in future research. In our personal experience, sleepiness typically sets in after around 1.5 h (Galley & Andrés, 1996).Whereas perceived driving errors cause the subject to scare up and result in temporary re-activation, the awareness of sleepiness alone is insufficient to cause sustained re-activation.

4.2.5 Interindividual differences. Substantial interindividual differences have been noted several times throughout our analyses. Obviously, not all subjects reported all stages of subjective fatigue. The stage of absolute alertness (10) was only reported by 14%, the stage of complete sleepiness (1) by 50%. We tried to point out this problem to the reader by showing the proportion of subjects that reported the corresponding alertness level in figures 4 to 8 and by initially analysing correlations on an individual level. Averaged individual correlations (column 2 and 4 in table 1) showed higher values than one global correlation across all subjects (column 6 in table 1). The 25-75th percentile ranges in figures 4 to 8 illustrate large variation of individual means or medians compared to global assessment. Furthermore, the other 50% of our subjects had even lower or higher values below and above the depicted 50%-range. This makes difficult the use of global thresholds to identify critical sleepiness in all subjects. For some parameters we even found 10-40% *paradoxical individual correlations* i.e. correlations indicating a relation contrary to group trends. As these correlations were significant in many cases and not related to restricted alertness ranges, they can hardly be dismissed as statistical artefacts or the well-known noise found in recording and automated processing of biosignals like the EOG. We rate the inaccuracy due to recording and automated processing at around 5-10%. Regarding paradoxical correlations, however, we assume that the three different processes constituting sleepiness while driving (decreasing attentional inhibition, increasing deactivation and self-activation) have subject-specific intensities. Therefore, in some drivers the attempts to counteract increasing sleepiness may arise earlier than in others. Reporting a strict distinction that for example changes in blink frequency indicate only light fatigue, and other parameters change exclusively during severe sleepiness for all subjects (Hargutt, 2003) may be the result of experiments with small sample sizes. Looking at 129 subjects of whom the majority passed through a broad range of self-rated alertness 8-2 (see figures 4-8, on average 6 steps per subject), we found that the identified processes change gradually in general, but show subject-specific courses with decreasing alertness.

4.2.6 Shortcomings of this study After discussing the results in detail, it is advisable to consider shortcomings of the present experimental setup and data treatment. The more or less linear decrease of self-rated fatigue with increasing time (see figure 2) has been observed frequently before (Reyner & Horne 1997) and this motivated us to interpolate between two ratings to be able to analyse oculomotoric events on an individual level. In this regard, we only address the homeostasis component of sleep models, i.e. increasing sleepiness with increasing time-since-sleep (Borberly & Achermann, 1999). Of course, interpolating linearly might raise linear correlations with these values which is why we also listed correlations with non-interpolated values (see table 1, column 6). The ranking of indicative oculomotoric parameters remains generally the same, speaking for reliability of these findings. However, as a consequence of our analyses we would recommend a more frequent asking of subjective alertness (at around every 20 minutes (see 4.2.1)) in order to obtain more data points.

The more fundamental question is whether alertness also decreases comparably in a real-life setting (Rizzo et al., 2007). Philip *et al.* (2005) conclude that the changes are elevated in a low-fidelity simulator setting, but are generally comparable to real-life settings. However, Belz *et al.*(2004) found neither driving parameters nor subjective ratings to be indicative of upcoming critical incidents in a real-life setting. In this latter study, the users were prompted by an alarm-tone to rate current fatigue on the Karolinska Sleepiness Scale while riding their trucks, and as the authors admit, this warning tone had a significant effect on video-rated fatigue, i.e. awakening the drivers. Thus, an electronic rating scale that lights up to prompt for a rating might be less obtrusive.

Although we do not know which version of an electronic KSS was used by Dingus and colleagues in their research, the reader may refer to the original studies (Dingus, Klauer et al., 2006; Dingus, Neale et al., 2006) for a more complete analysis of driving errors in relation to fatigue. Noteworthy is their *100-car Naturalistic Driving Study* (Dingus, Klauer et al., 2006), where the impact of fatigue was compared to all other kinds of inattention in a way that detailed what we were not able to come up with in our research with its focus on oculomotoric parameters.

4.2.7 Implications for future studies and development of drowsiness detectors. Finally, we would like to give some suggestion for further research and developments. In doing so we try to follow the stages of development outlined in the introduction, i.e. from general hardware and algorithmic considerations to aspects of a specific device.

First of all there is the question of reliable sensors for in-car measurement of eye- and lid-movements. As already stated in the introduction, the EOG with its self-adhesive electrodes will surely not be part of any future driver assistance system (Wright & McGown, 2001). We used it nevertheless to base our analysis of eye movement parameters on an established signal and we would generally recommend to also record a well-proven signal like the EOG when working with new sensors as for example Rimini-Doering *et al.* (2005) or Papadelis *et al.* (2007) did. Although there are attempts to include the sensors in glasses (Caffier et al., 2003; Johns et al., 2005; Johns et al., 2007), a remote camera is presumably favoured as it is less obtrusive. Appropriate camera-based sensors are being continuously improved, but they are still struggling with head turns, changes in ambient light and subjects wearing glasses when used under field conditions. These issues are amplified by the fact that most systems are provided as 'black boxes', i.e. researchers rarely have the opportunity to examine the underlying image processing procedures and must completely rely on the processed output. As development is usually done in parallel, any changes in the underlying image-processing or blink-detection algorithm require refinements of the sleepiness-detection algorithm that operates on that signal. Here, both sides might benefit from closer cooperation. Our broad listing of variables that can be obtained from eye movement recordings in turn could serve as a guide for deciding what aspects one should focus on. If a drowsiness detector should be capable of deploying saccadic in addition to blink parameters to reveal subjects that stare and rarely show blinks any more, then a substantial increase in sampling rate of the according sensors is unavoidable. The reliability of blink parameters like delay of lid reopening and lid closure speed would also profit from an increase in sampling rate.

Development of drowsiness-detection-algorithms (Thorslund, 2003; Svensson, 2004) should not confine itself too early to just blink duration and frequency, particularly if they are based on only a few experiments with small sample sizes. The individual differences we exemplified here are not restricted to possible thresholds indicating severe sleepiness, but also reflect individual differences in the actual indicative oculomotoric parameters. That is not to deny that a warning at blink durations larger than, for example, 500 ms is always justified, as the driver is not able to monitor the traffic adequately during that time and in most cases the blink in question will be a microsleep. However, the warning would be too late in our opinion, as microsleeps mostly appear during stages of severe sleepiness and can hardly be predicted as single events, but occur almost randomly during that stage (see figure 10 with blinks before and after a microsleeps). Fitness to drive will already be impaired much earlier, and microsleeps have to be considered as the *final evidence* rather than first hints of fatigue-related impairments. A drowsiness detector should account for that and warn prior to first appearances of microsleeps. A similar point could be made for lane departures: although this study was not aimed at analysing driving performance measures in detail and consequently the following remarks have to be taken with caution, we observed occasionally that drivers who showed almost no facial tonus any more and rated themselves very sleepy were still able to keep the lane. We are not sure whether an assistance system based on lane deviation alone could intervene on time to prevent a complete breakdown here. To what extent this phenomenon is limited to simulator settings or may even be pronounced in real life has to be examined in further research. However, we emphasize again that a warning system that is restricted to a distinct behavioural event might be too limited.

With regard to oculomotoric physiology, a three-stage model would make sense as many parameters show continuous changes across several alertness steps. It would be immediately intuitive when implemented together with a traffic light metaphor (green, yellow, red). Broader stages could also better compensate for temporary data loss or short-term attempts to deceive by the driver. A possible mapping of subjective alertness to traffic light stages would be:

- red: subjective alertness < 4
- yellow: subjective alertness 4-6
- green: subjective alertness > 6

The assumption of a three-stage model is also backed up by recent findings from EEG research (Trejo et al., 2007)

Individualised versions of a warning algorithm, concomitant to considering multiple oculomotoric parameters, may be unavoidable. This would be necessary to incorporate individual variability concerning thresholds as well as to include the most sensitive parameters. Depending on the respective user group, individualisation could be realized in two ways: On the one hand one could use a calibration run in the simulator, where an initially awake driver subjects himself to stages of critical fatigue, permitting a subsequent identification of the most indicative parameters. For professional drivers that use mostly the same vehicle such a procedure is justifiable, as a well-suited warner could ensure fitness to drive continuously (Milanovic & Klemenjak, 1999) and protect them and their cargo or passengers from serious harm.

On the other hand individualisation could be achieved by periodically reporting the current sleepiness assessment to the driver. Starting from relative changes in the standard variables blink duration, delay of reopening and blink interval the actual alertness stage could be feed back to the driver, who in turn would confirm or adjust the estimation through his steering-wheel-controls. His feedback will then be used to improve the prediction and to gradually determine the most indicative parameters. In general, drivers are aware of increasing sleepiness (Horne & Baulk, 2004; Nordbakke & Sagberg, 2007) and subjective alertness turned out to be a reliable measure in our experiments. Even when subjects systematically overrate their own alertness, microsleeps could still be registered as overlong lid closures. This would lead to warnings that would underscore a critical discrepancy between subjective and objective sleepiness. Besides improved sensitivity, such an approach would also emphasize the driver's personal responsibility and would prevent any illusion of an infallible warning device based on blink parameters, which is in our opinion not feasible at the moment.

Acknowledgements

Part of this research and the video ratings were done at the Center of Human-Machine-Systems (ZMMS), Berlin University of Technology.

We thank Dr. Volker Hargutt for several suggestions on an earlier version of our manuscript and Dr. Ron Mucha (both University of Würzburg) for addressing the language structure of our manuscript and for making many helpful comments.

Appendix A: standardisation formulas

General idea: standardised value = 100*(*actual value / value expected for this amplitude*)

Saccadic parameters

$\textit{standardised saccadic duration} = 100 * (\textit{saccadic duration}/(\textit{amplitude} * 2.07 + 26))$

$\textit{standardised saccadic speed} = 100 * (\textit{speed}/(445.9 * (1 - \exp(-0.04844 * \textit{amplitude} - 0.1121))))$

$\textit{standardised saccadic maximum speed} = 100 * (\textit{max. speed}/(580.4 * (1 - \exp(-0.06771 * \textit{amplitude} - 0.1498))))$

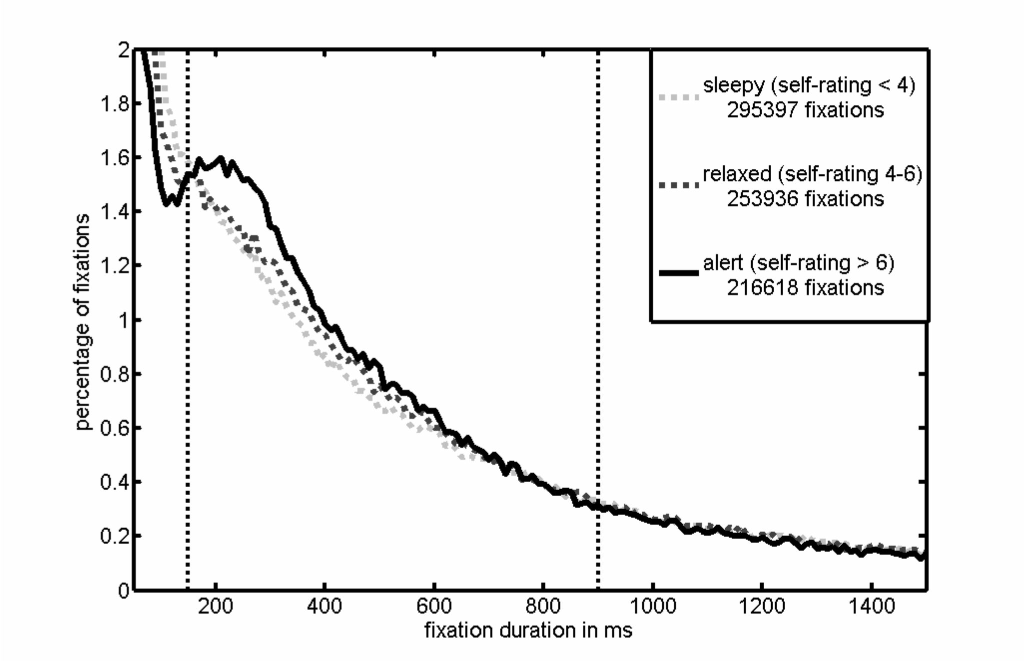
Blink parameters

$\textit{standardised blink duration} = 100 * (\textit{duration}/(0.862 * \textit{amplitude} + 121))$

$\textit{standardised lid closure speed} = 100 * (\textit{speed}/(7.186 * \textit{amplitude} + 60.55))$

$\textit{standardised maximum lid closure speed} = 100 * (\textit{max. speed}/(14.05 * \textit{amplitude} + 84.73))$

Appendix B: distribution of fixation durations



Distribution of fixations from 50 to 1500 ms during different stages of self-rated alertness. Vertical dotted lines indicate the range from 150-900 ms labelled as 'cognitive' saccades in our study. Note the high number of fixations around 200-400 ms in the alert state (solid black line) that dissappears with increasing fatigue. For further information on distribution of fixations see section 1.2.2 and Velichkovsky *et al.* (2001).

References

Abel, L. A.,& Hertle, R. W. (1988). Effects of psychoactive drugs on ocular motor behavior. In F. J. Pirozzolo (Ed.), *Neuropsychology of Eye Movements* (pp. 81-113). Hillsdale,N.J.: Erlbaum.

Adams-Guppy, J.,& Guppy, A. (2003). Truck driver fatigue risk assessment and management: a multinational survey. *Ergonomics, 46*, 763-779.

Ahsberg, E., Gamberale, F.,& Gustafsson, K. (2000). Perceived fatigue after mental work: an experimental evaluation of a fatigue inventory. *Ergonomics, 43*(2), 252-268.

Akerstedt, T.,& Folkard, S. (1994). Prediction of intentional and unintentional sleep onset. In J. R. Harsh (Ed.), *Sleep onset. Normal and abnormal processes* (pp. 73-87). Washington DC: American Psychological Association.

Annett, J. (2002). Subjective rating scales: science or art? *Ergonomics, 45*(14), 966-987.

App, E.,& Debus, G. (1998). Saccadic velocity and activation: development of a diagnostic tool for assessing energy regulation. *Ergonomics, 40*, 689-697.

Balkin, T. J., Bliese, P. D., Belenky, G., Sing, H., Thorne, D. R., Thomas, M., Redmond, D. P., Russo, M.,& Wesensten, N. J. (2004). Comparative utility of instruments for monitoring sleepiness-related performance decrements in the operational environment. *Journal of Sleep Research, 13*, 219-227.

Baulk, S. D., Reyner, L. A.,& Horne, J. A. (2001). Driver sleepiness-evaluation of reaction time measurement as a secondary task. *Sleep, 24*, 695-698.

Becker, W. (1989). Metrics. In M. E. Goldberg (Ed.), *The neurobiology of saccadic eye movements* (pp. 13-67). Amsterdam: Elsevier.

Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., Russo, M. B.,& Balkin, T. J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *Journal of Sleep Research, 12*, 1-12.

Belz, S. M., Robinson, G. S.,& Casali, J. G. (2004). Temporal separation and self-rating of alertness as indicators of driver fatigue in commercial motor vehicle operators. *Human Factors, 46*(1), 154-169.

Blatter, K., Graw, P., Munch, M., Knoblauch, V., Wirz-Justice, A.,& Cajochen, C. (2006). Gender and age differences in psychomotor vigilance performance under differential sleep pressure conditions. *Behavioural Brain Research, 168*(2), 312-317.

Blom, M. W., Bartel, P. R., de Sommers, K., Van der Meyden, C. H.,& Becker, P. J. (1990). The effects of alprazolam, quazepam and diazepam on saccadic eye movements, parameters of psychomotor function and the EEG. *Fundamental and Clinical Pharmacology, 4*, 653-661.

Borberly, A. A.,& Achermann, P. (1999). Sleep Homeostasis and Models of Sleep Regulation. *Journal of Biological Rhythms, 14*(6), 557-568.

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.

Cajochen, C., Brunner, D. P., Kräuchi, K., Graw, P.,& Wirz-Justice, A. (1995). Power density in theta/alpha frequencies of the waking EEG progressively increases during sustained wakefulness. *Sleep, 18*, 890-894.

Campagne, A., Pebayle, T.,& Muzet, A. (2004). Correlation between driving errors and vigilance level: influence of the driver's age. *Physiology & Behaviour, 80*(4), 515-524.

Collewijn, H., Erkelens, C. J.,& Steinman, R. M. (1988). Binocular co-ordination of human horizontal saccadic eye movements. *Journal of Physiology, 404*, 157-182.

Dinges, D. F.,& Kribbs, N. B. (1991). Performing while sleepy: effects of experimentally induced sleepiness. In T. Monk (Ed.), *Sleep, sleepiness, and performance* (pp. 97-128). Chichester: Wiley.

Dinges, D. F., Mallis, M. M., Maislin, G.,& Powell, J. W. (1998). *Final Report: Evaluation of Techniques for Ocular Measurement as an Index of Fatigue and as the Basis for Alertness Management* (DOT HS 808 762). Washington, D.C: National Highway Traffic Safety Administration.

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J.,& Knipling, R. R. (2006). *The 100-Car Naturalistic Driving Study: Phase II - Results of the 100-car field experiment* (DOT HS 810 593). Washington, DC: National Highway Traffic Safety Administration.

Dingus, T. A., Neale, V. L., Klauer, S. G., Petersen, A. D.,& Carroll, R. J. (2006). The development of a naturalistic data collection system to perform critical incident analysis: an investigation of safety and fatigue issues in long-haul trucking. *Accident Analysis and Prevention, 38*(6), 1127-1136.

Dorrian, J., Roach, G. D., Fletcher, A.,& Dawson, D. (2007). Simulated train driving: fatigue, self-awareness and cognitive disengagement. *Applied Ergonomics, 38*(2), 155-166.

Findlay, J. M.,& Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral & Brain Sciences, 22*, 661-674.

Galley, N. (1989). Saccadic eye movement velocity as an indicator of (de)activation. A review and some speculations. *Journal of Psychophysiology, 3*, 229-244.

Galley, N. (1993). The evaluation of the electrooculogram as a psychophysiological measuring instrument in the driver study of driver behaviour. *Ergonomics, 36*, 1063-1070.

Galley, N. (1998). An enquiry into the relationship between activation and performance using saccadic eye movement parameters. *Ergonomics, 40*, 698-720.

Galley, N.,& Andrés, G. (1996). Saccadic eye movements and blinks during long term driving on the autobahn with minimal alcohol ingestion. In S. P. Taylor (Ed.), *Vision in Vehicles V* (pp. 381-388). Amsterdam: Elsevier.

Galley, N.,& Churan, J. (2002). Ermüdung des Fahrers aus der Sicht des zukünftigen Fahrzeugs. In J. Kubitzki (Ed.), *Der sichere Fahrer. Ein Mythos. Erreichtes und Strittiges auf dem Gebiet der Verkehrssicherheit.* (pp. 99-113). Köln: TÜV-Verlag.

Galley, N., Schleicher, R.,& Galley, L. (in press). Blink parameters as indicators of driver's sleepiness - possibilities and limitations. In A. Gale (Ed.), *Vision in Vehicles X.*

Gevins, A., Leong, H., Dur, R., Smith, M. E., Le, J., DuRousseau, D., Zhang, J.,& Libove, J. (1995). Toward measurement of brain function in operational environment. *Biological Psychology, 40*, 169-186.

Gillberg, M., Kecklund, G.,& Akerstedt, T. (1996). Sleepiness and performance of professional drivers in a truck simulator - comparison between day and night driving. *Journal of Sleep Research, 5*, 12-15.

Godijn, R.,& Theeuwes, J. (2002). Programming of endogenous and exogenous saccades. Evidence for a competitive integration model. *Journal of Experimental Psychology. Human Perception and Performance, 28*, 1039-1054.

Hagenmeyer, L., van den Hurk, P., Nikolaou, P.,& Bekiaris, E. (2007). Development and Application of a Universal, Multimodal Hypovigilance-Management-System. In C. Stephanidis (Ed.), *Universal Access in HCI, Part II, HCII 2007, LNCS 4555,* (pp. 373-382). Berlin: Springer Verlag.

Häkkänen, H., Summala, H., Partinen, M., Tiihonen, M.,& Silvo, J. (1999). Blink duration as an indicator of driver sleepiness in professional bus drivers. *Sleep, 22*, 798-802.

Hargutt, V. (2003). *Das Lidschlagverhalten als Indikator für Aufmerksamkeits- und Müdigkeitsprozesse bei Arbeitshandlungen*. Düsseldorf: VDI Verlag.

Harrison, Y.,& Horne, J. A. (1996). Occurrence of "microsleeps' during daytime sleep onset in normal subjects. *Electroencephalography and Clinical Neurophysiology, 98*, 411-416.

Hasan, J.,& Broughton, R. (1994). Quantitative topographic EEG mapping during drowsiness at sleep onset. In J. R. Harsh (Ed.), *Sleep onset. Normal and abnormal processes* (pp. 219-235). Washington DC: American Psychological Association.

Hockey, G. R. J. (1993). Cognitive-energetical mechanisms in the management of work demands and psychological health. In L. Weiskrantz (Ed.), *Attention, selection, awareness and control* (pp. 328-354). Oxford: Oxford University Press.

Hoffman, A. C. (1946). Eye movements during prolonged reading. *Journal of Experimental Psychology, 36*, 95-118 Zs A 2255.

Hofmann, M. (2005). *eogui - ein Computerprogramm zur automatischen Lidschlag- und Sakkadenerkennung in Matlab.* Diplomarbeit, FH Schmalkalden, Schmalkalden.

Hofmeister, J., Heller, D.,& Radach, R. (1998). The return sweep in reading. In T. Mergner (Ed.), *Current oculomotor research: physiological and psychological aspects* (pp. 349-357). New York: Plenum Press.

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

Horne, J. A.,& Reyner, L. A. (1995). Driver sleepiness. *Journal of Sleep Research, 4*, 23-39.

Ingre, M., Akerstedt, 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.

Ingre, M., Akerstedt, T., Peters, B., Anund, A., Kecklund, G.,& Pickles, A. (2006). Subjective sleepiness and accident risk avoiding the ecological fallacy. *Journal of Sleep Research, 15*, 142-148.

Johns, M., Tucker, A.,& Chapman, R. (2005). *Monitoring the Drowsiness of Drivers: A New Method Based on the Velocity of Eyelid Movements.* Paper presented at the International ITS Conference, San Francisco, November 2005.

Johns, M. W., Tucker, A., Chapman, R., Crowley, K.,& Michael, N. (2007). Monitoring eye and eyelid movements by infrared reflectance oculography to measure drowsiness in drivers. *Somnologie - Schlafforschung und Schlafmedizin*.

Kaida, K., Akerstedt, 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*(6), 1324-1331.

Karrer, K., Briest, S., Vöhringer-Kuhnt, T., Baumgarten, T.,& Schleicher, R. (2005). Driving without awareness. In G. Underwood (Ed.), *Traffic and transport psychology. Proceedings of the ICTTP 2004* (pp. 455-470). Amsterdam: Elsevier.

Kribbs, N. B.,& Dinges, D. (1994). Vigilance decrement and sleepiness. In J. R. Harsh (Ed.), *Sleep onset. Normal and abnormal processes* (pp. 113-125). Washington DC: American Psychological Association.

Lal, S. K.,& Craig, A. (2002). Driver fatigue: electroencephalography and psychological assessment. *Psychophysiology, 39*, 313-321.

Lal, S. K. L.,& Craig, A. (2001a). A critical review of the psychophysiology of driver fatigue. *Biological Psychology, 55*, 173-194.

Lal, S. K. L.,& Craig, A. (2001b). Electroencephalography activity associated with driver fatigue: implications for a fatigue countermeasure device. *Journal of Psychophysiology, 15*, 183-189.

Lavie, P. (1991). The 24-hour sleep propensity function (SPF): practical and theoretical implications. In T. Monk (Ed.), *Sleep, sleepiness, and performance* (pp. 65-96). Chichester: Wiley.

Lavie, P. (2001). Sleep-wake as a biological rhythm. *Annual Revue of Psychology, 52*, 277-303 Zs A 1743.

Lavine, R. A., Sibert, J. L., Gokturk, M.,& Dickens, B. (2002). Eye-tracking measures and human performance in a vigilance task. *Aviation, Space, and Environmental Medicine, 73*, 367-372.

Luckiesh, M.,& Moss, F. K. (1937). The eyelid reflex as a criterion of ocular fatigue. *Journal of Experimental Psychology, 20*, 589-608.

Mallis, M. M., Mejdal, S., Nguyen, T. T.,& Dinges, D. F. (2004). Summary of the Key Features of Seven Biomathematical Models of Human Fatigue and Performance. *Aviation, Space, and Environmental Medicine, 75*, A4-14.

Matthews, G.,& Desmond, P. A. (2002). Task-induced fatigue states and simulated driving performance. *Quarterly Journal Experimental Psychology A, 55*, 659-686.

Meinold, P. (2005). *Psychologie des Lidschlags. Eine literatur- und methodenkritische Studie.* Philosophische Fakultät, Universität zu Köln,

Milanovic, M.,& Klemenjak, W. (1999). *Psychologische Maßnahmen gegen ermüdungsbedingte Unfälle bei Fernfahrern.* Wien: Kuratorium für Verkehrssicherheit.

Milosevic, S. (1997). Drivers' fatigue studies. *Ergonomics, 40*, 381-389.

Moller, H. J., Kayumov, L.,& Shapiro, C. M. (2003). Microsleep episodes, attention lapses and circadian variation in psychomotor performance in a driving simulation paradigm. In D. V. McGehee, J. D. Lee, M. Rizzo, M. Raby & L. Noyle (Eds.), *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 130-137). Iowa City, Iowa: University of Iowa.

Morris, T. C.,& Miller, J. C. (1996). Electrooculographic and performance indices of fatigue during simulated flight. *Biological Psychology, 42*, 343-360.

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

Ogilvie, R. D., Wilkinson, R. T.,& Allison, S. (1989). The detection of sleep onset: behavioral, physiological, and subjective convergence. *Sleep, 12*, 458-474.

O'Hanlon, J. F.,& Kelley, G. R. (1977). Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicle operation. In R. R. Mackie (Ed.), *Vigilance* (pp. 189-202). New York: Plenum Press.

Ongerboer de Visser, B. W.,& Bour, L. J. (2006). Eye and eyelid movements during blinking: an eye blink centre? *Supplements Clinical Neurophysiology, 58*, 16-25.

Pannasch, S., Dornhoefer, S. M., Unema, P. J. A.,& Velichkovsky, B. M. (2001). The omnipresent prolongation of visual fixations: saccades are inhibited by changes in situation and in subjects's activity. *Vision Research, 41*, 3345-3351.

Papadelis, C., Chen, Z., Kourtidou-Papadeli, C., Bamidis, P. D., Chouvarda, I., Bekiaris, E.,& Maglaveras, N. (2007). Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. *Clinical Neurophysiology, 118*(9), 1906-1922.

Philip, P., Sagaspe, P., Taillard, J., Moore, N., Guilleminault, C., Sanchez-Ortuno, M., Akerstedt, T.,& Bioulac, B. (2003). Fatigue, sleep restriction, and performance in automobile drivers: a controlled study in a natural environment. *Sleep, 26*(3), 277-280.

Philip, P., Sagaspe, P., Taillard, J., Valtat, C., Moore, N., Akerstedt, T., Charles, A.,& Bioulac, B. (2005). Fatigue, sleepiness, and performance in simulated versus real driving conditions. *Sleep, 28*(12), 1511-1516.

Philip, P., Taillard, J., Klein, E., Sagaspe, P., Charles, A., Davies, W. L., Guilleminault, C.,& Bioulac, B. (2003). Effect of fatigue on performance measured by a driving simulator in automobile drivers. *Journal of Psychosomatic Research, 55*(3), 197-200.

Pinel, J. P. J. (1997). *Biopsychologie* (4th.ed. ed.). Heidelberg: Spektrum.

Pivik, R. T. (1991). The several qualities of sleepiness: psychophysiological considerations. In T. Monk (Ed.), *Sleep, sleepiness, and performance* (pp. 3-38). Chichester: Wiley.

Porcu, S., Bellatreccia, A., Ferrara, M.,& Casagrande, M. (1998). Sleepiness, alertness and performance during a laboratory simulation of an acute shift of the wake-sleep cycle. *Ergonomics, 41*(8), 1192-1202.

Radach, R., Heller, D.,& Inhoff, A. (1998). Occurence and function of very short fixation durations in reading. In T. Mergner (Ed.), *Current oculomotor research: physiological and psychological aspects* (pp. 321-332). New York: Plenum Press.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Reitter, C. (2001). *Beanspruchung des Autofahrers durch Monotonie- eine Untersuchung mit Blick- und lidmotorischen Parametern.* Berlin: Dissertation an der TU Berlin.

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

Rimini-Doering, M., Altmueller, T., Ladstaetter, U.,& Rossmeier, M. (2005). Effects of Lane Departure Warning on Drowsy Drivers' Performance and State in a Simulator. In L. Boyle, J. D. Lee, D. V. McGehee, M. Raby & M. Rizzo (Eds.), *Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 88-95). Iowa City, Iowa: University of Iowa.

Rizzo, M., Robinson, S.,& Neale, V. L. (2007). The Brain in the Wild - Tracking Human Behavior in Natural and Naturalistic Settings. In R. Parasuramaran & M. Rizzo (Eds.), *Neuroergonomics*. New York City, NY: Oxford University Press.

Saito, S. (1992). Does fatigue exist in a quantitative measure of eye movements? *Ergonomics, 35*, 607-615.

Sirevaag, E. J.,& Stern, J. A. (2000). Ocular measures of fatigue and cognitive factors. In R. W. Backs & W. Boucsein (Eds.), *Engineering Psychophysiology* (pp. 269-288). Mahwah, NJ: Lawrence Erlbaum Associates.

Skipper, J. H.,& Wierwille, W. W. (1986). Drowsy driver detection using discriminant analysis. *Human Factors, 28*, 527-540.

Smit, A. C.,& Van-Gisbergen, J. A. (1990). An analysis of curvature in fast and slow human saccades. *Experimental Brain Research, 81*, 335-345.

Stern, J. A., Boyer, D.,& Schroeder, D. (1994). Blink rate: a possible measure of fatigue. *Human Factors, 36*, 285-297.

Summala, H., Häkkänen, H., Mikkola, T.,& Sinkkonen, J. (1999). Task effects on fatigue symptoms in overnight driving. *Ergonomics, 42*, 798-806.

Svensson, U. (2004). *Blink behaviour based drowsiness detection - method development and validation.* Biomedical Engineering, University Linköping, Linköping.

Tassi, P.,& Muzet, A. (2001). Defining the states of consciousness. *Neuroscience & Biobehavioral Reviews, 25*, 175-191.

Thackray, R. I. (1981). The stress of boredom and monotony: a consideration of the evidence. *Psychosomatic Medicine, 43*, 165-176.

Thiffault, P.,& Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis and Prevention, 35*(3), 381-391.

Thorslund, B. (2003). *Electrooculogram analysis and development of a system for defining stages of drowsiness.* Biomedical Engineering, Linköpings Universitet, Linköping.

Trappenberg, T. P., Dorris, M. D., Munoz, D. P.,& Klein, R. M. (2001). A model of saccade initiation based on the competitive integration of exogeneous and endogeneous signals in the superior colliculus. *Journal of Cognitive Neuroscience, 13*, 256-271.

Trejo, L. J., Knuth, K., Prado, R., Rosipal, R., Kubitz, K., Kochavi, R., Matthews, B.,& Zhang, Y. (2007). EEG-Based Estimation of Mental Fatigue: Convergent Evidence for a Three-State Model. In C. Stephanidis (Ed.), *Universal Access in HCI, Part II, HCII 2007, LNCS 4555* (pp. 201-211). Berlin: Springer Verlag.

Tucker, A.,& Johns, M. W. (2005). The Duration of Eyelid Movements During Blinks: Changes with Drowsiness. *Sleep, 28*, (Suppl) A122.

Unema, P., Dornhoeffer, S., Steudel, S.,& Velichkovsky, B. (2000). An attentive look at driver's fixation durations. In e. al. (Ed.), *Vision in Vehicles VII*. Amsterdam: North Holland.

Van Dongen, H. P. A., Baynard, M. D., Maislin, G.,& Dinges, D. F. (2004). Systematic interindividual differences in neurobehavioral impairment from sleep loss: evidence of triat-like differential vulnerability. *Sleep, 27*, 423-433.

Van Dongen, H. P. A., Vitellaro, K. M.,& Dinges, D. F. (2005). Individual differences in adult human sleep and wakefulnes: leitmotif for a research agenda. *Sleep, 28*, 479-496.

van Orden, K. F., Jung, T. P.,& Makeig, S. (2000). Combined eye activity measures accurately estimate changes in sustained visual task performance. *Biological Psychology, 52*, 221-240.

Velichkovsky, B. M., Dornhoefer, S. M., Pannasch, S.,& Unema, P. J. A. (2001). Visual fixations and level of attentional processing. In A. Duhowski (Ed.), *Proceedings of the International Conference of Eye Tracking Research & Applications*. Palm Beach Gardens, FL: ACM Press.

Velichkovsky, B. M., Joos, M., Helmert, J. R.,& Pannasch, S. (2005). *Two visual systems and their eye movements: evidence from static and dynamic scene perception.* Paper presented at the CogSci 2005: Proceedings of the XXVII Conference of the Cognitive Science Society. July 21-23 Stresa, Italy, pp. 2283-2288., July 21-23 Stresa, Italy.

Verwey, W. B.,& Zaidel, D. M. (2000). Predicting drowsiness accidents from personal attributes, eye blinks and ongoing visual behavior. *Personality and Individual Differences, 28*, 123-142.

Victor, T. W., Harbluk, J. L.,& Engström, J. A. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F, 8*, 167-190.

Vöhringer-Kuhnt, T., Baumgarten, T., Karrer, K.,& Briest, S. (2004). *Wierwille's method of driver drowsiness evaluation revisited.* Paper presented at the 3rd International Conference on Traffic and Transport Psychology. Nottingham, UK.

Webb, W. B. (1994). Prediction of sleep onset. In J. R. Harsh (Ed.), *Sleep onset. Normal and abnormal processes* (pp. 53-72). Washington DC: American Psychological Association.

Wierwille, W. W. (1999). Historical perspective on slow eylid closure: Whence PERCLOS? In R. J. Carroll (Ed.), *Ocular measures of driver alertness: Technical conference procedings* (pp. 31-53). Washington DC: Federal Highway Administration.

Wierwille, W. W.,& Ellsworth, L. A. (1994). Evaluation of driver drowsiness by trained raters. *Accident Analysis and Prevention, 27*, 571-581.

Wright, K. P., Badia, P.,& Wauquier, A. (1995). Topographical and temporal patterns of brain activity during transition from wakefulness to sleep. *Sleep, 18*, 880-889.

Wright, N.,& McGown, A. (2001). Vigilance on the civil flight deck: incidence of sleepiness and sleep during long-haul flights and associated changes in physiological parameters. *Ergonomics, 44*(1), 82-106.