

MICROSLEEP EPISODES AND RELATED CRASHES DURING OVERNIGHT DRIVING SIMULATIONS

Martin Golz^{1,5}, David Sommer¹, Jarek Krajewski², Udo Trutschel^{3,5}, & Dave Edwards⁴

¹Faculty of Computer Science, University of Applied Sciences Schmalkalden, Germany

²Work and Organizational Psychology, University of Wuppertal, Germany

³Circadian Technologies, Inc., Stoneham, Massachusetts, USA

⁴Machine Research, Caterpillar, Inc., Peoria, Illinois, USA

⁵Institute for System Analysis and Applied Numerics, Tabarz, Germany

E-mail: m.golz@fh-sm.de

Summary: Microsleep (MS) episodes and related crashes were studied in an overnight driving simulation study. A new definition of MS proposed recently was applied and the mean number as well as the mean length of MS was calculated. MS occurred much more frequently than crashes. Within all pre-crash intervals (length 1 minute) the percentage of MS was calculated. Results showed that there are numerous MS episodes before every crash. The mean length of MS was between 5 and 9 seconds and did not change significantly during the night. The mean MS percentage was high within pre-crash intervals (60-80%) and is a predictor for crashes.

INTRODUCTION

Two important consequences of sleepiness during driving simulations are microsleep (MS) and crashes. The rate of both events increases if sleepiness as well as monotony is increasing. To keep monotony on a high level, the simulated traffic complexity should be low, no kind of communication should be allowed and a non-animating environment should be presented. Overnight driving in the fully dark lab with sparsely illuminated driving scenes fulfills this condition. In order to increase sleepiness, the time since sleep and the time on driving task should be long. In addition, the time of day should be selected such that the driver passes the circadian trough. The latter factor is very important for drivers having high circadian amplitudes, such as young drivers.

Unfortunately, the combination of all conditions is sometimes fulfilled on the real road (e.g. overnight highway driving after an active day and evening). The combination is also important for studies in the lab in order to get enough examples to drive statistical investigations. In this paper we present statistical results on MS events and crashes and ask how they are related in time. How many MS events occur in the average immediately before a crash happened? How many crashes occur without any MS events beforehand? How is their temporal development across the night? And how is the inter-individual variation?

The definition of MS comprises a problem. On the one hand several authors have defined MS strictly by EEG criteria no matter what behavior is observed simultaneously. Then again, MS can be defined in behavioral terms at the cost of strictness because of many different behavioral signs with large inter-individual variations. In addition, some of them are vague to assess. Another difficulty is that sometimes subjects display signs of oculomotoric quiescence or blank stare.

A third way which combines both types of criteria has been proposed by (Sommer, Golz, Schnupp, Krajewski, Trutschel, Edwards, 2009). In a first step only events with clear behavioral signs were considered, such as prolonged eyelid closures, slow roving eye movements, head noddings, slow drifting head movements. All of them are often followed by abrupt reactions. The occurrence of these behavioral signs led to the class label 'MS' for an adaptive biosignal processing (EEG, EOG) and pattern recognition methodology. EEG/EOG segments accompanied by behavioral signs of high sleepiness, but still active driving and fulfilling the lane tracking were labeled as the other class ('Non-MS'). This way, behavioral criteria and biosignal criteria were combined by modern, automatically learning computer algorithms (Golz and Sommer, 2008). It has been demonstrated that the MS detector gained high sensitivity as well as high specificity to strictly defined behavioral MS. Furthermore, many events were detected in the EEG/EOG which are similar to behavioral MS, but no strict behavioral signs of MS occurred simultaneously (Sommer, 2009; Sommer, Golz, Schnupp, Krajewski, Trutschel, Edwards, 2009). The detected MS events correlated strongly with independent subjective self-ratings of sleepiness as well as objective measures of driving performance, e.g. standard deviations of lateral position in lane. In this paper, this combined definition of EEG/EOG-based detection of MS which are related to evident behavioral events is utilized.

METHOD

Driving simulations were conducted at the driving simulation lab of the University of Applied Sciences, Schmalkalden, Germany. The study was designed to investigate driving performance and subjects' behaviour under high level of monotony and sleepiness. Monotony was supported by selecting very low traffic density (no car in lane, 1 car every 3 minutes in the opponent lane; road configuration: winding two-lane road, undulating landscape). Subjects were instructed to keep in lane as best as possible and to avoid falling asleep. After returning from MS events subjects were reminded that if driving performance becomes too bad or signs of behavioral MS returned too fast, the experiment would be terminated.

Crashes have been defined as intervals where all four wheels were out of lane, no matter if the car went to the left or to the right. Incidents, such as 1, 2, or 3 wheels out of lane were not regarded as crashes. When crashes appeared, an extensive soundscape, as well as video scene, was played to increase the emotional importance of this event, to set a short break in the run of the driving simulation and to warn the driver by the observers.

Subjects

Fourteen healthy young volunteers ranging from 18-32 years of age (mean 22.4; SD 4.1) and had held a driver's licence for at least 1 year. 1 male and 2 female quitted driving because of simulator sickness, 1 male quitted because of back pain.

Driving Simulations

Ten subjects (8 male, 2 female) completed all 7 overnight driving runs starting at the top of every hour (1:00–7:00 AM). Each run had a duration of 40 minutes and was preceded and followed by vigilance tests and response-to-sleepiness questionnaires. Reports of vigilance tests will be given

elsewhere. Time since sleep was at least 16 hours, checked by actigraphy. Subjects have been prepared beforehand by at least one hour training on the simulator.

Measurements

Several biosignals were recorded: EEG (F1, F2, C3, Cz, C4, O1, O2, A1, A2, com.av.ref.), EOG (vertical, horizontal), ECG, EMG (m. submentalis). In addition, three video recordings (driver's head & pose, driver's eyes, driving scene) were stored. Also several variables of the car (e.g. time series of steering angle and lateral position of the vehicle) were sampled.

Subjectively experienced sleepiness was rated every 4 minutes during driving following suggestions of (Åkerstedt et al., 2006). Subject's response was given orally using the Karolinska Sleepiness Scale (KSS) (Åkerstedt, 1990). Further experimental details have been published elsewhere (Golz et al., 2007).

Visual Ratings of MS

Two operators who watched the video streams performed a first judgment of ongoing MS immediately during the experiments. Typical signs of MS are prolonged eyelid closures, slow roving eye movements, head noddings, slow drifting head movements, and major driving incidents. They were often followed by abrupt reactions. Several other signs were observed, but it has been decided not to solely rely on them. Some examples are bursts of alpha and theta activity in the EEG, spontaneous pupil contractions and blank stare. In all, we have found 2,290 MS events (per subject: mean number 229 ± 67 , range 138–363).

For the detection of MS events on a second by second basis a careful determination of the point in time where MS is starting will be needed. Therefore, all recorded video material and biosignals underwent off-line scoring made by an independent and trained rater. He refined and eventually corrected the results of online ratings.

Visual ratings were only utilized to select examples of MS and Non-MS out of the continuum of the recordings. This leads to so-called data labeling needed for supervised learning algorithms like Support-Vector Machines (see below).

Signal Analysis

Segments of EEG/EOG during MS were processed by a multi-stage classification methodology. Segment length should range between 4 and 12 seconds (Golz et al., 2007). We used 6s and calculated spectral power densities by the modified periodogram method. Afterwards logarithmic scaling and summation in narrow spectral bands (width 1 Hz) in the range 0.5 to 23 Hz turned out to be important to reduce the error rate in the subsequent classification step. Support-Vector Machines with Gaussian kernel function were applied to these data in order to perform classification (i.e. to separate signal segments labeled as MS from segments labeled as Non-MS). After a series of parameter optimizations, an accurate MS detector was constructed. The mean accuracy of 97.7% has been estimated on evident examples of behavioral MS (Sommer et al. 2009). In a second step, the MS detector has been applied consecutively to all data (not only

evident MS and Non-MS examples) to detect further examples which are similar in the EEG/EOG characteristics.

RESULTS

With increasing time since sleep the mean number of detected MS events as well as the number of crashes increased (Figure 1). MS occurred much more frequently than crashes. In the first four driving runs (1:00–4:00 AM) almost no crash happened. Afterwards, rapidly increasing crash probabilities were observed. Several subjects were not longer able to perform the driving task. From the beginning of the night, the mean number of detected MS was high and increased by factor 3 to the end of the night.

The length of MS is a highly varying quantity, both inter- and intra-individually. Their mean value ranged between 5 and 9 seconds and increased slightly with time since sleep (Figure 1).

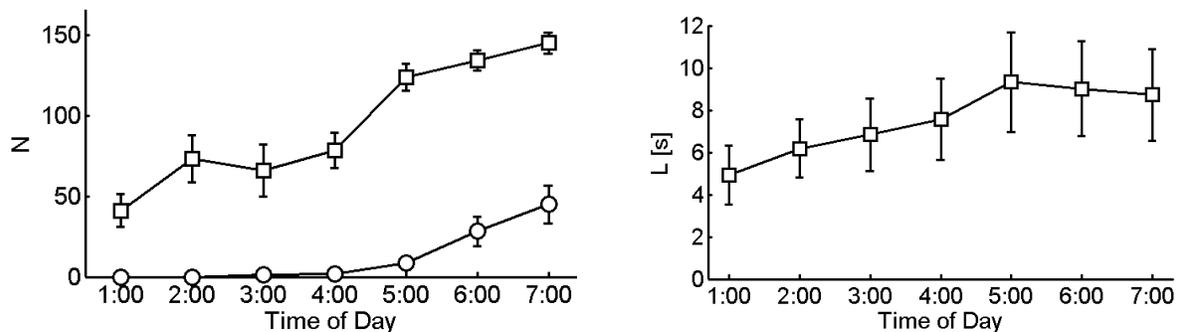


Figure 1. Results of MS detection: The mean number of MS (left; squares), of crashes (left; circles), and the mean length of MS (right) increase with time of day
(Data were averaged across subjects; the bars represent standard errors)

To answer the question if and how many MS periods occurred preceding crashes, a fixed interval prior to each crash was analysed. The length was chosen to 60s and it was estimated how much of the time MS was detected. The resulting quantity is called MS percentage and is 100% if during all 60s prior to a crash MS was detected. If no MS was detected during 60s prior to a crash, then a MS percentage of 0% resulted. Both extremes didn't occur. Since crashes occurred almost always in the second half of the night, where both number and length of MS were high, the percentage of MS prior to crashes was relatively high (Figure 2). The mean values varied between subjects in the range between 55% and 78%. The number of crashes varied highly between subjects (Figure 2, upper line). One subject performed with less than 10 crashes whereas other subjects had much more than 100 crashes over the whole night.

Immediately during the seconds from crashes high MS percentages occurred. Short intervals completely free of MS constituted an exception. Their frequency rapidly decreased with interval length (Figure 2). Only less than 2% of all crashes happened with no MS during pre-crash intervals when interval length was greater than 5s. In the present study no crash was found without MS during 180s long pre-crash intervals.

DISCUSSION

This pilot study demonstrated the application of a new way to assess MS episodes with high temporal resolution. Short as well as long MS episodes were detected. The mean length of MS was between 5 and 9 seconds and was slightly but not significantly dependent on time since sleep. Results suggest that crashes are signaled by MS episodes, although an accurate prediction of crashes based on MS was not possible. There were many MS episodes which were not followed by a crash.

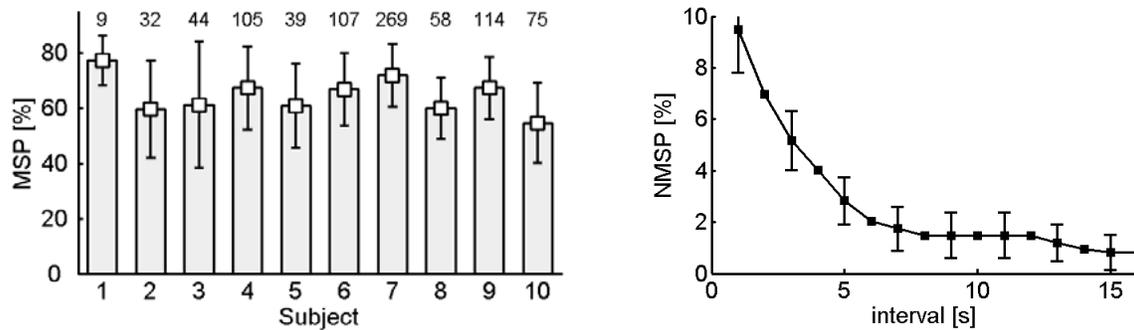


Figure 2. Results of crash analysis: Mean and standard deviations of MS percentage (MSP) during the 1 min interval preceding every crash; Mean MSP ranges between 55% and 78% (left); The percentage of pre-crash intervals containing no MS is below 10% and decreases rapidly with increasing length of the interval (right)
(Data were averaged across subjects; the bars represent standard errors)

The quantification of the EEG/EOG suspected to be MS during pre-crash intervals, which we called MS percentage, is a new measure (Sommer, Golz, Schnupp, Krajewski, Trutschel, Edwards, 2009). The consecutive application of an MS detection algorithm on the entire EEG/EOG led to a binary information (MS or Non-MS) with high temporal resolution (100ms). Averaging across intervals (length 1 min) led to this new variable. It was demonstrated that the EEG/EOG is highly susceptible to MS if analysis was done at least 5s prior to crashes.

Driving performance during MS episodes was also investigated in terms of steering variability, variation of lateral position in lane, minimum time to lane crossing (Paul, Boyle, Tippin, Rizzo, 2005). It has been shown that there is a remarkable change compared to both intervals pre- and post-MS (interval length 3s). The mean variation of the lateral position in lane almost doubled during MS and the mean minimum time to lane crossing was reduced by ca. 40%. Therefore, in their study a remarkable increase of crashes in succession to MS must be assumed.

Lower effects of MS on driving performance variables were reported recently (Boyle, Tippin, Paul, Rizzo, 2008). They demonstrated large differences between straight and curved roadway segments. Non-MS versus MS episodes were characterized by insignificant changes in the mean speed (< 1%), in the steering entropy (< 3%), and in the standard deviation of the slow wave EEG activity (+2.2% on straight roadways, +10.4% in curves). Both measures of lane tracking performance, the standard deviation of position in lane and the minimum time to lane crossing were sensitive to MS, but depend highly on the roadway type. The first variable increased by +12.7% and +23.7%, and the second variable increased by +24.2% and decreased by -23.2% on straight and curved roadway segments, respectively. Crashes were not considered.

Risser, Ware, Freeman (2000) found that MS defined as EEG episodes of increased alpha or theta activity lasting more than 3s correlated with lane position variability and crash frequency. Crashes occurred always in later sessions and were relatively rare events, whereas MS episodes were significantly more frequent, which was also reported by other authors for daytime driving simulations (Moller, Kayumov, Bulmash, Nhan, Shapiro, 2006). They investigated EEG defined MS episodes as a consequence of daytime sleepiness in healthy normals. A much lower mean number of MS episodes was reported ranging between 0.6 and 1.2. Three main reasons may explain why our result was much higher (ranging between 40 and 140). Firstly, the authors' definition of MS was completely different. EEG activity lasting 3-30s were scored if alpha, theta activity, or true sleep EEG dominated. Our MS detector methodology oftentimes found shorter events. Moreover, EEG with more complex patterns were classified to MS episodes, because evident behavioral MS also occurred with such EEG characteristics (Golz, Sommer, Chen, Trutschel, Mandic, 2007). Secondly, our study protocol provided partial sleep deprivation. Time since sleep was between 17 and 23 hours as against 2 and 8 hours in their daytime experiments. Thirdly, our driving sessions lasted 10 minutes longer. This is important, because time on task effects are known to play a role on MS. Subsequent analysis showed that the mean number of MS episodes reduces by 34.1% if driving simulations would have been reduced from 40 minutes to 30 minutes and the number of crashes reduces by 43.0%.

Nevertheless, both, the results of (Moller, Kayumov, Bulmash, Nhan, Shapiro, 2006) as well as the present study demonstrate that the number of MS episodes across a single driving session and crash risk correlated highly. Future investigations should ask if the incidence of MS episodes and their temporal evolution are useful as an indicator of upcoming crashes. Visual inspection of our data suggested that this relation is complex and not easily recognizable. It may be that a prediction of crashes is not possible. For now it can be stated that in sleep deprivation studies the crash risk is highly associated with numerous MS suspected EEG/EOG immediately before.

ACKNOWLEDGMENT

This study was funded by the Federal Ministry of Education and Research within the research program "Research at University of Applied Sciences together with Enterprises" under the project 176X08.

REFERENCES

- Boyle, L.N., Tippin, J., Paul, A., & Rizzo, M. (2008). Driver Performance in the Moments Surrounding a Microsleep. *J Transport Res Part F*, 11, 126-136.
- Davidson, P.R., Jones, R.D., & Peiris, M.T.R. (2007). EEG-Based Lapse Detection with High Temporal Resolution, *IEEE Trans Biomed Engin*, 54(5), 832-839.
- Golz, M., Sommer, D., Trutschel, U., Sirois, B., & Edwards, D. (2010). Evaluation of Fatigue Monitoring Technologies. *J Somnology*, 14(3), 187-189.
- Golz, M., & Sommer, D. (2008). Automatic knowledge extraction: fusion of human expert ratings and biosignal features for fatigue monitoring applications. In D. Mandic, M. Golz, A. Kuh, D. Obradovic & Tanaka T. (Eds.) *Signal processing techniques for knowledge extraction and information fusion*. New York: Springer.

- Golz, M., Sommer, D., Chen, M., Trutschel, U., & Mandic, D. (2007). Feature Fusion for the Detection of Microsleep Events. *J VLSI Signal Proc Syst*, 49, 329-342.
- Moller, H.J., Kayumov, L., Bulmash, E.L., Nhan, J., & Shapiro, C.M. (2006). Simulator performance, microsleep episodes, and subjective sleepiness: normative data using convergent methodologies to assess driver drowsiness. *J Psychosomatic Res*, 61, 335-342.
- Paul, A., Boyle, L.N., Tippin, J., & Rizzo, M. (2005). Variability of Driving Performance During Microsleeps, *Proc 3rd Int Symp Driver Assessment*, Rockport, Maine, 18-24.
- Risser, M.R., Ware, J.C., & Freeman, F.G. (2000). Driving Simulation with EEG Monitoring in Normal and Obstructive Sleep Apnea Patients. *Sleep*, 23(3), 393-398.
- Sommer, D. (2009). *Microsleep characterisation utilizing neuroinformatic methods*. PhD Thesis (in German). Ilmenau University of Technology, Ilmenau, Germany.
- Sommer, D., Golz, M., Schnupp, T., Krajewski, J., Trutschel, U., & Edwards, D. (2009). A Measure of Strong Driver Fatigue. *Proc 5th Int Symp Driver Assessment*, Big Sky, Montana, 9-15.