Use of Subjective and Physiological Indicators of Sleepiness to Predict Performance during a Vigilance Task

Kosuke KAIDA1–3*, Torbjörn ÅKERSTEDT2,3, Göran KECLUND2,3, Jens P. NILSSON3 and John AXELSSON2,3

1 National Institute of Occupational Safety and Health, Nagao 6–21–1, Tama-ku, Kawasaki 214-8585, Japan
2 Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden
3 National Institute for Psychosocial Medicine (IPM), Stockholm, Sweden

Received January 23, 2007 and accepted April 18, 2007

Abstract: Sleepiness is a major risk factor for serious injury and death in accidents. Although it is important to develop countermeasures to sleepiness to reduce risks, it is equally important to determine the most effective timing for these countermeasures. To determine optimum timing for necessary countermeasures, we must be able to predict performance errors. This study examined the predictability of subjective and physiological indicators of sleepiness during a vigilance task. Thirteen healthy male volunteers (mean age, 26.9 yr; SD = 5.98 yr; range 22–43 yr) participated in the study. Participants used the Karolinska sleepiness scale (KSS) to rate their subjective sleepiness every 4 min during a 40-min Mackworth clock test. Electrophysiological and performance data were divided into 10 epochs (i.e., 1 epoch lasted for 4 min). To estimate predictability, the data from the sleepiness indicators used for the correlation analysis were preceded by one epoch to the performance data. Results showed that sleepiness indicators (KSS score and electroencephalographic [EEG] alpha activity) and standard deviation of heart rate (SDNN) were significantly correlated with succeeding performance on the vigilance test. These findings suggest that the KSS score, EEG alpha activity, and SDNN could be used to predict performance errors.

Key words: Prediction, Sleepiness, EEG, Heart rate variability, Karolinska sleepiness scale

Introduction

Sleepiness is a major risk factor for serious injury and death in accidents1–5). Effective countermeasures against sleepiness have been examined such as caffeine6), short nap7–10), and bright light exposure11–13). Although the development of sleepiness countermeasures is important, it is equally important to determine the most effective timing for applying these countermeasures. This can be done by predicting performance errors.

The prediction of performance errors should be made from both long-term and short-term perspectives. For long-term predictions, the circasemidian rhythm of sleepiness12) is a general predictor of performance errors (e.g., sleepiness and the risk for accidents increase in the early morning and in the afternoon13). For short-term predictions, indicators such as subjective sleepiness score, electroencephalographic (EEG) activity while the eyes are open (alpha [8.0–12.0 Hz] and theta [4.0–7.9 Hz])13–16), and eye-blinks17–21) can be used. Another indicator for short-term prediction would be heart rate variability (HRV) which is believed to be an indicator of autonomic nervous activity22, 23). From previous studies it is known that high cognitive performance is correlated with the high frequency component (HF) of HRV24) and the root mean square difference among successive R-R normal intervals (rMSSD)25). A rather high correlation was also
observed between HF and rMSSD, suggesting that parasympathetic nervous activity (or vagal activity)26 relates to cognitive performance. However, few studies have examined whether sympathetic nervous activity, which can be estimated by low frequency component (LF) and standard deviation of heart rate (SDNN), correlates to performance.

In the present study, we focused on short-term prediction of performance errors and examined whether sleepiness indicators and HRV predict performance errors during a vigilance task. The experimental context was set similar to general work situations (i.e., a monotonous control room), and attention was sustained during moderate sleep deprivation in the experiment, as this is a common requirement for many shift workers. The analysis was performed using data from a study on how the act of rating subjective sleepiness influences other indicators of sleepiness and its masking effect on performance27.

Methods

Participants and design
Thirteen healthy male volunteers (mean, 26.9 yr; SD, 5.98 yr; range, 22–43 yr) participated. All participants met the following criteria: (1) no physical or mental health problems, (2) no shift work within 3 months before the experiment, (3) no travel to a different time zone within 3 months before the experiment, (4) no medications, (5) nonsmoker, (6) body mass index (BMI) <25 (calculated as weight in kg divided by the square of the height in meters). The average BMI of the participants was 22.8 (SD = 2.48) kg/m².

The participants took part in two experimental conditions during the same day, one with repeated sleepiness ratings (rating) and one without (non-rating). The orders of the conditions were counterbalanced in the participants. In the rating condition, participants rated their subjective sleepiness every 4 min during 40 min while performing a task. Only data from the rating condition was used for the analysis in the present study as subjective rating was one of the variables to be examined.

The experimental protocol was reviewed and approved by the regional ethical committee of Stockholm (2005/412-31/1), using the Helsinki committee rules. The experiment was carried out during May–July 2005.

Procedure
A week before the experimental day, participants came to the laboratory and gave written informed consent after the procedures were fully explained. They then practiced the performance task (described below). Participants’ sleep-wake cycles were monitored by an Actiwatch® (Cambridge Neurotechnology, Cambridge, UK) and they filled out a sleep diary for 7 days before the experiment. Participants were asked to abstain from beverages containing caffeine and alcohol on the experimental day.

All participants were instructed to sleep between 03.00 h and 07.00 h on the night before the experiment to induce sleepiness. To facilitate adherence, participants were asked to send an e-mail or an SMS (Short Message Service) to the experimenter every hour using their mobile phone from 23.00 h until their scheduled bedtime (03.00 h). The wake-up time (07.00 h) was confirmed by a phone call from the participants to the experimenter. Adherence was also monitored by the Actiwatch28.

On the experimental day, participants arrived at the laboratory at 08.00 h. After breakfast (cereal, yogurt juice, and toast, served by the experimenter), electrodes were placed on the participant. From 09.30 h, the Karolinska drowsiness test (KDT), which consists of 5 min with the eyes open and 5 min with the eyes closed29, was carried out with the participant lying on a bed in a supine position to check the initial sleepiness/arousal level.

A 40-min performance task was started at 09.50 h and finished at 10.30 h (session 1). After a 10-min break, a second session of the same task was performed from 10.40 h to 11.20 h (session 2). The rating condition was assigned for session 1 or 2 with counterbalance. During the rating condition, the experimenters entered the room and asked the participant to rate his subjective sleepiness verbally using the modified version of the Karolinska sleepiness scale (KSS)29 while the vigilance task was still running. The use of the scale had been practiced beforehand, and consists of the following scores: 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 awake, 8 = sleepy, some effort to keep awake, 9 = very sleepy, great effort to keep awake, struggling against sleep, 10 = extremely sleepy, falls asleep all the time. The KSS, which has been used in a number of studies, responds sensitively to sleep loss30, 31 and is closely related to performance and physiological indicators of sleepiness14, 16, 29. The experiment was carried out in a temperature- and light-controlled (21–23°C, <70 lux) and sound-attenuated room. The result of the KDT is presented elsewhere27.

Performance task
The performance task was a modified version of the Mackworth clock test32. This test measures the ability to sustain attention in the face of monotonous stimulation.
Participants were sitting 80 cm away from a screen on which the stimuli were presented. On the screen, a filled black circle appears for 0.6 s at one of 14 positions arranged in a circle, creating the impression of a dot that jumps in a circular pattern across the screen. The interval between jumps is 0.6 s. During the 40-min test, one of the positions was randomly skipped 60 times. The task was to push a button whenever the one position was skipped. From the task, the number of misses (i.e., omissions) was calculated. A response was considered as an omission/miss if the reaction time was longer than 1,800 ms. In the Mackworth clock test, participants are asked to detect an absence of stimulus instead of detecting an appeared stimulus. In such a task, it is hard to determine stimulus onset and participants need to concentrate their attention to the accuracy, not the reaction time, to improve their performance. Thus we chose accuracy as a dependent variable. To obtain a normal distribution, the performance data were transformed as sqrt (x) + sqrt (x+1) where “x” is the obtained result.

**Recordings**

EEG electrodes (Ag/AgCl) were attached to the scalp at the C3 referenced to A2. For blink detection using the electro-oculogram (EOG), the electrodes were positioned immediately above the eyebrow, above the center of the left eye and at 1 cm lateral to the outer canthus of the left eye. Electrodes were also attached below the chin for a bipolar submental electromyogram (EMG). Heart rate was derived from cardiac intervals based on continuous electrocardiographic (ECG) recordings from a modified lead 1 (subclavical) placement. The sampling rate for all channels was 200 Hz (16-bit analog-to-digital conversion) and the low pass filter time constants were 0.3 s for the EEG and the ECG, 3.2 s for the EOG, and 0.03 s for the EMG. The low pass filter was set at 30 Hz for the EEG and EOG, and at 60 Hz for the ECG. Electrode impedance was maintained below 5 kΩ.

**Spectral power analysis of the EEG**

Alpha (8.0–12.0 Hz) and theta (4.0–7.9 Hz) activity were calculated using the fast Fourier transform with a Hamming window. Power spectra were calculated for every 4 s of EEG data. The adjacent 4-s values were averaged to yield power density values for 4 min. Artifacts in the EEG were removed by visual inspection before the spectrum calculation. The spectral analysis was performed using Somnologica software version 3.2.1 (Medcare Flaga Medical Devices, Reykjavik, Iceland).

**Heart rate variability**

Heart rate variability was calculated using ectopic-free ECG recordings in 4-min intervals for time domain analysis. The ECG signal was fed into a generator that produced a pulse at the rising phase of each R wave. The standard deviation of normal R-R intervals (SDNN) and the root mean square difference among successive R-R normal intervals (rMSSD) were calculated. SDNN and rMSSD are commonly used to evaluate heart rate variability and they are considered as a measure of autonomic activity.

The 12th order model of Burg autoregressive spectral estimation (parametric method) was performed as a frequency domain analysis. Samples for the spectral analysis were fixed at 200 consecutive ectopic-free heart beats. The mean heart rate was 64.4 beats per minute (SD = 8.81), and the analysis epoch was around 3.1 min (SD = 0.51) (depending on the time required to collect 200 heart beats). The heart rate sampling was always started from the starting point of every 4-min interval, as was the time domain analysis. The powers in the low-frequency (LF) band (0.04–0.15 Hz) and in the high-frequency (HF) band (0.15–0.50 Hz) were calculated. The LF-to-HF ratio (LF/HF) was calculated using the power in each band. This involved the R-R spectral power components: HF power, reflecting parasympathetic activity; LF power, reflecting a predominance of sympathetic activity with a parasympathetic component; and the LF/HF, regarded as an index of sympathovagal balance. All the analyses of HRV were performed using Somnologica software version 3.2.1.

**Statistical analysis**

The electrophysiological and performance data were divided into 10 epochs in each task (i.e., 1 epoch lasted 4 min) because the KSS was scored every 4 min. To estimate
the correlation and predictability of the sleepiness indicators for performance, a correlation analysis was performed between sleepiness indicators and the performance errors of each individual. The data from the sleepiness indicators and HRV used for the correlation analysis were preceded by one epoch to the performance data (i.e., the data of the sleepiness and HRV indicators in a certain epoch were analyzed against the data of performance in the immediate succeeding epoch, i.e., “prediction”). For instance, the correlation analysis was performed between the sleepiness indicators in epoch 1 and performance in epoch 2, between the indicators in epoch 2 and performance in epoch 3, and so on. Thus, the study examined whether the performance errors would be predicted by the sleepiness indicators recorded in the preceding epoch (4 min). The difference between “correlation” and “prediction” is explained in Fig. 1. Pearson product-moment correlation coefficients (r) were averaged across individuals and tested for significance using a one-sample t-test (p<0.05).

To analyze change across time, indicators which correlate to performance error were submitted to a one-way repeated-measures analysis of variance (ANOVA). To control for the type 1 error associated with violation of the sphericity assumption, degrees of freedom greater than one were reduced by the Huynh-Feldt epsilon (ε) correction. All statistical analyses were performed with the SPSS program (SPSS, Chicago, IL).

Results

Total sleep/rest time estimated by the Actiwatch for the night before the experimental day was 271.8 min (SD = 39.02). The cardiac data of one participant was excluded from the analysis of session 1 due to a large variation from the group mean beyond standard deviations of 2.5. Mean values, SD, and the range of the original data are shown in Table 1.

The “correlations” and “predictions” (correlations between sleepiness indicators and succeeding performance, that is, “lagged” correlations) are shown in Table 2. The number of omissions (misses) was correlated to KSS score, EEG alpha activity, SDNN of HRV. The number of omissions was predicted by increased KSS score, EEG alpha activity, SDNN, and rMSSD. Other indicators such as the number of long eye-blinks did not show any significant correlation and prediction. The ranges of the correlation coefficients (r) and the number of negative correlations or predictions are presented in Table 2 to show individual differences.

Performance errors (omissions) \[ F(9, 108) = 4.48, p<0.01, \ ε = 1.00 \], KSS score \[ F(9, 108) = 4.00, p<0.01, \ ε = 1.00 \], alpha power density \[ F(9, 108) = 10.35, p<0.01, \ ε = 0.76 \], and SDNN \[ F(9, 108) = 17.07, p<0.01, \ ε = 0.65 \] significantly increased as time elapsed. However, rMSSD did not change as time elapsed \[ F(9, 108) = 0.84, p = 0.58, \ ε = 1.00 \].
show data fluctuations and compare the indicators, data was translated to z-score and depicted in Fig. 2.

Discussion

The present study showed that the number of omissions on the vigilance test could be predicted by the KSS score, EEG alpha activity, SDNN and rMSSD. Although the non-lagged correlations (i.e., “correlations”) between the KSS score, EEG activity, and performance errors have been reported in previous studies, the present study went a step further and predicted performance errors from preceding subjective and physiological data.

In particular, the performance prediction by SDNN is an interesting finding since no negative “prediction” was found in this indicator, while individual differences of “prediction” were observed in most indicators (See Table 2). This suggests that SDNN increases monotonically and less influenced by individual differences, predicting the increase of succeeding performance errors. It has been known that people in a drowsy situation repeat sleep onset and arousal, in which heart rate decreases with sleep onset and increases with arousal. The results of the present study suggest that heart rate might fluctuate when the participants are struggling with severe sleepiness and the fluctuation might be reflected in the time domain of HRV.

Although the “predictions” examined in this study were statistically significant, standard deviations of correlation efficiency (r) were larger than averaged data in some indicators. However, the standard deviation in SDNN was relatively small. This suggests that SDNN (and alpha power density) may be a useful performance predictor. Since heart rate can be relatively easily obtained, it could be a useful indicator in the development of a system to predict performance errors for drivers or people working in other

<table>
<thead>
<tr>
<th>Variable</th>
<th>“Correlation”</th>
<th>“Prediction”</th>
</tr>
</thead>
<tbody>
<tr>
<td>MISS</td>
<td>0.03 (0.36)</td>
<td>0.33 / 0.82</td>
</tr>
<tr>
<td>SDNN</td>
<td>0.16 (0.21)</td>
<td>0.33 (0.17)</td>
</tr>
<tr>
<td>rMSSD</td>
<td>0.16 (0.21)</td>
<td>0.33 (0.17)</td>
</tr>
<tr>
<td>LF</td>
<td>0.11 (0.34)</td>
<td>0.08 (0.37)</td>
</tr>
<tr>
<td>HF</td>
<td>0.06 (0.40)</td>
<td>0.11 (0.32)</td>
</tr>
<tr>
<td>HR</td>
<td>-0.03 (0.35)</td>
<td>0.08 (0.28)</td>
</tr>
<tr>
<td>Theta</td>
<td>0.13 (0.30)</td>
<td>0.20 (0.38)</td>
</tr>
<tr>
<td>KSS</td>
<td>0.23 (0.33)</td>
<td>0.22 (0.37)</td>
</tr>
<tr>
<td>EEG (C3) Alpha</td>
<td>0.25 (0.35)</td>
<td>0.36 (0.29)</td>
</tr>
<tr>
<td>EEG (C3) Theta</td>
<td>0.13 (0.30)</td>
<td>0.20 (0.38)</td>
</tr>
</tbody>
</table>

Bold type = significant at p<0.05.

“Correlation” means normal correlations among variables and “prediction” means correlations of variables and the succeeding performance.

MISS, the number of omission.
Negative, the number of participants who showed negative prediction.

Fig. 2. Time course of the indicators.

z-scores were transformed in each individual as (x–x) / SD where “x” is the obtained result.

Industrial Health 2007, 45, 520–526
INTERINDICATORS OF SLEEPINESS TO PREDICT PERFORMANCE ERRORS

risky environments.

In previous studies, it was demonstrated that increased HRV related to improved performance. Hansen et al. (2003) reported that a group of people with high rMSSD showed better performance than those with low rMSSD. Kohler et al. (2006) showed that caffeine increased rMSSD. These results suggest that rMSSD and cognitive performance are positively correlated. In contrast, rMSSD demonstrated a significant “prediction” of decreased performance in the present study, that is, rMSSD and performance are negatively correlated. The previous studies measured trait HRV and performance. On the other hand, the present study examined state HRV variation during a drowsy situation, in which struggling with severe sleepiness might affect rMSSD. Additionally, there is a finding that rMSSD is predictive of accuracy especially in the complex cognitive task. The performance task used in the present study might have been too simple compared to the previous studies and it might have led to a different result. To clarify this, further investigations are necessary.

In the present study, HF and LF of HRV did not correlate to, nor predict, performance. Hansen et al. (2004), who examined trait HRV before and after physical training, reported that HF level increased after 4 week training and performance was also improved. In the frequency domain analysis of HR, a relatively stable change of heart rate (<0.05 Hz) was analyzed to estimate autonomic nervous activity. On the other hand, in the time domain analysis, short term fluctuation of HR would be detected. Thus, the time domain analysis might be better to detect dynamic fluctuations of HR for short term prediction of performance errors.

Apart from HRV, one study reported that the number of eye-blinks can be used to detect sleep onset and it is related to sleep latency on the multiple sleep latency test. In the present study, however, we could not find any “prediction” and “correlation” between performance and the number of eye-blinks. Obviously, blinking is affected by several factors such as task demand, mental and visual workload, and position of gaze. In addition, one might conceive of blinking as a strategy to counteract sleepiness although this does not seem to have been studied. Therefore, it seems unlikely that the number of blinking would be a reliable indicator (or predictor) of sleepiness and performance.

A limitation of the present study is that the 4-min time window was probably too wide for some of the measures. Some of the variables might have become significant if the epoch (i.e., “lag” for prediction) had been shorter. Although this limitation should be considered in further studies, in conclusion, the present study showed that time domain heart rate variability, EEG alpha activity, and the KSS score could be used to predict performance errors on a vigilance test.

Acknowledgements

This work was supported by the EU integrated project SENSATION (IST 507231). The principal author is the recipient of the Japan Society for the Promotion of Science (JSPS) postdoctoral fellowship.

This work was performed at the National Institute for Psychosocial Medicine (IPM), Stockholm, Sweden.

References


