J. Sleep Res. (2006) 15, 291–300

Frequent lapses of responsiveness during an extended visuomotor tracking task in non-sleep-deprived subjects

MALIK T. R. PEIRIS[1,2], RICHARD D. JONES[1,2,3,4], PAUL R. DAVIDSON[1,3], GRANT J. CARROLL[1,5] and PHILIP J. BONES[1,2]

[1]Van der Veer Institute for Parkinson's and Brain Research, [2]Electrical and Computer Engineering, University of Canterbury, [3]Medical Physics and Bioengineering, Christchurch Hospital, [4]Medicine, Christchurch School of Medical and Health Sciences, University of Otago and [5]Neurology, Christchurch Hospital, Christchurch, New Zealand

Accepted in revised form 26 June 2006; received 24 October 2005

SUMMARY We investigated the occurrence of lapses of responsiveness (lapses) in 15 non-sleep-deprived subjects performing a 1D continuous tracking task during normal working hours. Tracking behaviour, facial video, and electroencephalogram (EEG) were recorded simultaneously during two 1-h sessions. Rate and duration were estimated for lapses identified by a tracking flat spot and/or video sleep. Fourteen of the 15 subjects had one or more lapses, with an overall rate of 39.3 ± 12.9 lapses per hour (mean ± SE) and a lapse duration of 3.4 ± 0.5 s. We also found that subjects' performance improved towards the end of the 1-h long session, even though no external temporal cues were available. Spectral power was found to be higher during lapses in the delta, theta, and alpha bands, and lower in the beta, gamma, and higher bands, but correlations between changes in EEG power and lapses were low. In conclusion, lapses are a frequent phenomenon in normal subjects – even when not sleep-deprived – engaged in an extended monotonous continuous visuomotor task. This is of particular importance to the transport sector in which there is a need to maintain sustained attention for extended periods of time and in which lapses can lead to multiple-fatality accidents.

KEYWORDS behavioural microsleep, drowsiness, EEG power spectra, lapse, visuomotor tracking

INTRODUCTION

Excessive drowsiness at work is an important safety issue in modern societies and is usually associated with insufficient or poor-quality sleep because of irregular sleep patterns or sleep disorders. Nevertheless, even well-rested individuals who are not sleep-deprived may experience 'sleep-related states' without a preceding phase of subjectively experienced drowsiness (Sagberg, 1999). Drowsy individuals performing an extended active task, such as driving, often cycle rapidly between periods of wake and sleep, as exhibited by cyclical variation in both electroencephalogram (EEG) power spectra and task performance measures (Makeig *et al.*, 2000). Although well documented in sleep-restricted people, these episodes are less well described in rested non-sleep-deprived individuals.

Lapses of responsiveness (lapses) are brief episodes in which a subject unintentionally stops responding to the task they are performing. By contrast, the related term *microsleep* is usually used to describe brief episodes (min 1–15 s, max = 14–30 s) of EEG-defined sleep (Harrison and Horne, 1996; Hemmeter *et al.*, 1998; Priest *et al.*, 2001; Tirunahari *et al.*, 2003; Valley and Broughton, 1983). While many lapses are associated with EEG-defined microsleeps, and hence are apparently caused by a low arousal, the link between EEG-defined sleep and lapse behaviour is generally not strong (Ogilvie, 2001). Non-arousal-related lapses in sustained or task-directed attention also occur and may show quite distinct characteristics (Neale *et al.*, 2005; Parasuraman and Davies, 1984).

Lapses are of particular concern in occupations in which public safety depends on extended unimpaired performance, such as truck drivers, locomotive drivers, pilots, air traffic

Correspondence: Richard Jones, Van der Veer Institute for Parkinson's and Brain Research, 16 St Asaph Street, Christchurch, New Zealand. Tel.: +64-3-378-6077; fax: +64-3-378-6080; e-mail: richard.jones@vanderveer.org.nz

controllers, health professionals, and process control workers. Lapses in such occupations can have disastrous consequences, including multiple fatalities. For example, reduced driver alertness has been suggested as a cause of many train accidents, especially those occurring during the night (Torsvall and Åkerstedt, 1987). There is also increasing recognition that drowsiness leads to increased risk of accidents in private motor vehicle drivers (Cummings *et al*., 2001; Lal and Craig, 2001; Sagberg, 1999; Stutts *et al*., 2003). A substantial survey in the UK found that 29% of the participants considered themselves to have come close to falling asleep while driving in the previous 12 months (Maycock, 1997). Another study found that falling asleep was the likely cause of 16% of all accidents on major roads in south-west England and over 20% on midland motorways (Horne and Reyner, 1995).

Numerous groups have demonstrated lapses of performance under monotonous task conditions with and, less commonly, without sleep deprivation. A number of studies have reported lapses in sleep-deprived subjects performing a variety of auditory and visual sustained attention tasks, and mental tasks (Cajochen *et al*., 1999; Dinges *et al*., 1997; Doran *et al*., 2001; Kleitman, 1963; Schroeder *et al*., 1994; Torsvall and Åkerstedt, 1988). In addition, lapses have also been reported during simulated night-time driving (Arnedt *et al*., 2005). Bills (1931) reported an increase in the number of 'blocks' (or lapses) in non-sleep-deprived subjects with time-on-task in subjects performing a mental task. More recently, Van Orden *et al*. (2000) reported lapses in a group of non-sleep-deprived subjects performing a continuous 2D tracking task during normal waking hours.

Sleep deprivation has a strong influence on level of drowsiness and, consequently, task performance. Using a 10-min psychomotor vigilance task (PVT), Dinges *et al*. (1997) determined that the maximum lapse rate after 7 days of sleep restriction (5 h of sleep per night) in a group of 20 young adults was 24 lapses per hour (reaction time > 500 ms). Lapsing did not increase between the second baseline night and the first night of sleep restriction but increased after that. After 70 h of total sleep deprivation, Doran *et al*. (2001) measured a maximum PVT error rate (false start trials and anticipatory responses less than 100 ms) of 96 errors per hour. In a study by Torsvall and Åkerstedt (1988), participants 'dozed off' 7.4 times during a 45-min discrete visual vigilance task which commenced at 01:30 hours. Doran *et al*. (2001) suggested that performance during sleep deprivation is highly unstable because of the subjects rapidly fluctuating between states which cannot be defined as fully awake or asleep because of the influence of sleep-initiating mechanisms.

Time-on-task also plays a key role in performance fluctuations during mental tasks (Bills, 1931; Dinges *et al*., 1997; Thiffault and Bergeron, 2003). Mackworth (1969) suggested that decline in performance during a monotonous task may be because of habituation of the neural response because of repetitive stimuli.

While circadian factors have a strong effect, task and environmental conditions also influence task performance. Almost all studies examining driver fatigue suggest that it occurs more rapidly on monotonous roads such as highways (Thiffault and Bergeron, 2003). Subjects sitting in a comfortable chair in a quiet, dimly lit, and warm room, whom have been asked to refrain from making unnecessary movements, tend to lose vigilance and have intermittent lapses in task performance, even when well rested (Makeig and Jung, 1996; Valley and Broughton, 1983; Van Orden *et al*., 2000).

In this study, we investigated lapses occuring in non-sleep-deprived young adults during normal working hours. We also investigated a conservatively defined subcategory of lapses during which clear behavioural signs of sleep, such as eyelid closure and head-nodding, were evident. We call these *behavioural microsleeps* (BMs) to emphasize their arousal-related nature and to distinguish them from EEG-defined microsleeps.

In our study, subjects carried out a 1D continuous tracking task (CTT) while full-head EEG, eye movements, video, and tracking response kinematics were recorded. The focus of this paper is on describing the occurrence and characteristics of lapses and BMs in a group of subjects superficially considered unlikely to fall asleep during a task. We derived conservative estimates of BM rate and duration by searching for clear lapses of task responsiveness in tracking performance (indicated by the cessation of movement of the response cursor for an extended period while the target is moving – henceforth referred to as a *flat spot*) occurring concurrently with behavioural sleep observed on the video recordings (video BM). In addition to the conservative estimate, we also calculated rate and duration estimates for lapses based on the occurrence of either a flat spot in the tracking response or a video BM. Finally, we also investigated changes in EEG spectral power in the standard bands associated with lapses.

METHODS

Subjects

Fifteen normal healthy male volunteers (students and hospital staff), aged 18–36 years (mean = 26.5), were recruited for the study. No subject reported a current or previous neurological or sleep disorder and all had visual acuities of 6/9 (=20/30) or better in each eye. All subjects considered that they slept normally the previous night (mean = 7.8 h, SD = 1.2 h, min = 5.1 h) and, hence were considered non-sleep-deprived. Ethical approval for the study was obtained from the Canterbury Ethics Committee.

Apparatus

Subjects performed a 1D pursuit continuous visuomotor tracking task. The tracking system comprised a PC with two 17-inch color monitors, a Matrox G550 dual-head graphics card (Matrox Graphics Inc., Dorval, Quebec, Canada), and an Advantech PCL 1710 data acquisition board (Advantech Co. Ltd., Taipai, Taiwan). The tracking task was generated by the sensory-motor tests program *SMTests*™ (Jones *et al*., 1993) and a steering wheel (395 mm diameter, wheel-to-screen gain = 1.075 mm per degree) was used to control an arrow-shaped cursor, located at the bottom of the screen. The eye-to-screen distance was 136 cm. Subjects were provided with an 8-s preview of a pseudo-random target (bandwidth 0.164 Hz, period 128 s) which scrolled downwards at a rate of 21.8 mm s^{-1} (Jones, 2000). It was generated by summing 21 sinusoids with random phases. The task required smooth movements of the steering wheel over a 175° range and measured a subject's ability to keep the point of the arrow on the moving target. The position of the wheel was sampled at 64 Hz using a potentiometer mounted on the shaft of the wheel. An analogue video camera was used to record head and facial features of the subject during the session (25 Hz frame rate).

Using an in-house-developed EEG system, EEG was recorded from electrodes at 16 scalp locations, band-pass filtered (0.5–100 Hz), and digitized at 256 Hz with a 16-bit A/D converter. Bipolar derivations used to calculate power spectra were Fp1–F7, F7–T3, T3–T5, T5–O1, Fp2–F8, F8–T4, T4–T6, T6–O2, Fp1–F3, F3–C3, C3–P3, P3–O1, Fp2–F4, F4–C4, C4–P4, and P4–O2.

Procedure

All experimental sessions were held between 12:30 and 17:00 hours. Each subject attended two sessions held at least 1 week apart (mean 17 days, range 7–50). Subjects were asked to refrain from consumption of stimulants or depressants, such as caffeine, nicotine, alcohol, illicit drugs, and medications, for 4 h prior to each session; other than this, no constraint was placed on the content or timing of their pre-session lunch. Before performing the task, they were asked to keep the point of the arrow as close as possible to the target throughout the session. The time of day and remaining task time were not provided during the task. However, subjects were informed of the task duration (1 h) prior to the experimental session. Subjects were asked to stay alert and perform to the best of their ability and, aside from eye blinks, to keep their eyes open as much as possible during the task. They sat in a comfortable chair and the room temperature was maintained between 22 and 25 °C. At the end of the session each subject was asked whether they considered they had 'lapsed' at any point.

Flat spot detection

Lapses in tracking performance are most obvious when the response cursor simply stops moving for an extended period while the target is moving or when the tracking response is non-coherent with the target. Only the first category, which we refer to as *flat spots*, were included in our intentionally conservative analysis, as lapses in the second category are difficult to identify with confidence. Flat spots occurring when the target velocity is approximately zero (at turning points) were not counted, as at these times the subject can track adequately without moving the response cursor.

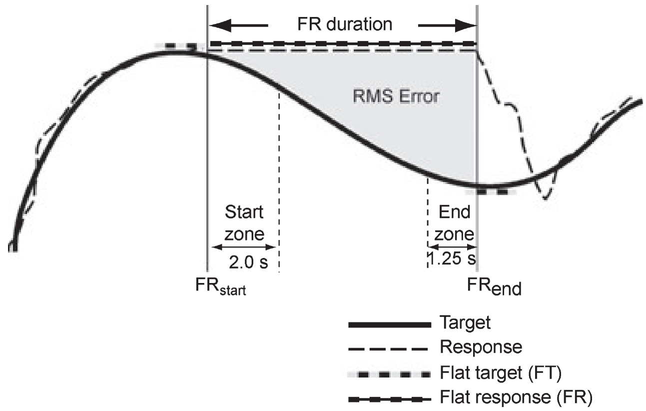

Figure 1. Flat spots were detected using flat segments of the target and response waveforms, RMS error between the target and the response during a flat response (FR), and the duration of a flat response (FR duration). Flat target (FT) sections are segments in the target waveform of at least 300 ms duration and with a within-segment deviation less than 1.5 mm. The 2.0-s start zone and the 1.25-s end zone are also marked within the FR (an FT can occur within either zone but not both).

A schematic of the procedure used to detect flat spots is shown in Fig. 1. The target and response signals were low-pass filtered with a cut-off at 5 Hz using an eighth-order bidirectional Butterworth filter. Target and response sections were then marked using the following criteria: (a) *flat target region* – a flat segment of the target, defined as a section of at least 300 ms duration, within which the deviation was ≤1.54 mm and (b) *flat response region* – a flat segment of the response, defined as a section of at least 1500 ms duration, within which the deviation was ≤0.77 mm. A 2.0-s 'start zone' and a 1.25-s 'end zone' were marked within each flat response region.

A flat response region was classified as a flat spot if (A) there were no flat targets within the 'start' or 'end' zones of the flat response, (B) only one flat target occurred within the 'start' or 'end' zones of the flat response, (C) the RMS error between the target and the response during the flat response region was greater than 15.0 mm or (D) the duration of the flat response region was ≥6.0 s. Criterion C was necessary for cases where a flat response coincided with two or more flat target regions, as such a flat response would not be detected as a flat spot using criterion A or B. However, by utilizing the large error between the target and response during a flat response, we were able to correctly identify many of these as flat spots. Visual inspection of the tracking data confirmed that an RMS error threshold of 15.0 mm was sufficiently sensitive to detect these types of flat spots without introducing false positives. The duration of the longest contiguous flat target region was 5.0 s, so any flat responses ≥6.0 s had to be flat spots.

Video rating

The video recording of each session was conservatively rated by one of the authors (MP), without knowledge of the corresponding tracking performance. Being 'blind' to tracking performance ensured that an independent measure of alertness was obtained from the video data. The video was rated on a six-level scale: 1 = alert, 2 = distracted, 3 = forced eye closure while alert (FEC), 4 = light drowsy, 5 = deep drowsy, and 6 = sleep (including microsleep). We used criteria similar to Wierwille and Ellsworth (1994) to define our video rating scale. Alert periods were identified by the presence of features such as fast eye blinks and normal facial tone in video recordings. Intervals when the subject appeared diverted from the task at hand were rated as 'distracted'. Instances of the subject intentionally closing their eyes while alert (typically to relieve eye fatigue) were marked as 'forced eye closure'. Instances of the subject appearing subdued, with slower blinks, were rated as 'light drowsy'. 'Deep drowsy' was identified by a paucity of eye movements, decrease of facial tone, and partial closure of the eye lids. Based on video alone, deciding between light and deep drowsy was found to be particularly difficult. Sleep events – *video BMs* – were identified by prolonged eye-lid closure, sometimes accompanied by rolling upward or sideways movements of the eyes, head-nodding, and often terminated by waking head jerks. Transitions in the video recording had a time resolution of 1.0 s.

Definite BM rate and duration

Intervals in which flat spots and video BMs overlap in time are referred to as *definite BMs*. To ensure our estimated event rates and durations were conservative, slightly different methods were used to calculate each. Definite BM duration was obtained by finding all samples for which a video BM and a flat spot occurred simultaneously (Fig. 2). To obtain a conservative estimate of definite BM rate, a logical OR was

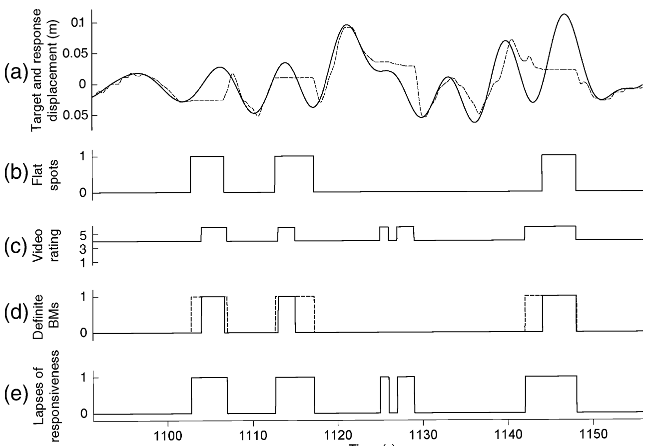

Figure 2. An illustrative example of the methodology used to obtain estimates of definite BM and lapse counts and durations. Data from a subject approximately 18 min into the 1-h tracking task: (a) tracking behavior including target (solid) and response (dotted), (b) detected flat spots (c) independent video rating of the subject's level of alertness, (d) definite BMs in terms of lower bounds for count (dashed) and duration (solid), and (e) lapses. There is little doubt that a lapse occurred between 1124 and 1129 s but, as no response flat spot was identified, no definite BM was identified. However, according to the video rating, two 'sleep' events occurred during the same interval and hence two lapses were identified.

applied between video BMs and flat spots, checking that both a video BM and a flat spot occurred during that period. This ensured that only a single event was counted if there were multiple flat spots during a video BM and vice versa. Except when calculating rate, the duration method was always used to identify definite BMs.

As the pseudo-random target had a period of 128 s, a bin width of 128 s was used to calculate event rate changes as this provided identically difficult target segments and enabled comparison of performance across epochs. Event rate was calculated in terms of the onset of events.

Lapse rate and duration

The presence of either a video BM *or* a flat spot (cf. *both* needed for a definite BM) provides a sound, if conservative, indicator of the presence of a lapse. The minimum detection duration for all events (flat spots, video BMs, definite BMs, and lapses) was 1.0 s. Fig. 2 illustrates how the lapse and definite BM rate and duration measures were derived.

EEG analysis

A notch filter was applied to remove 50 Hz interference from the EEG and independent components analysis was applied to remove eye blink artifacts (Delorme and Makeig, 2004; Jung *et al*., 2000). A window size of 512 samples (2 s) and an overlap of 50% between successive windows were used to calculate power spectra via a 40th-order autoregressive Burg model; a high model order was required to obtain adequate separation of the spectral bands of interest. Data in each window was detrended prior to the calculation of power spectra. A window overlap of 1 s allowed a temporal resolution of 1 s for the spectral power in the delta (1.0–4.5 Hz), theta (4.5–8.0 Hz), alpha (8.0–12.5 Hz), beta (12.5–25.0 Hz), gamma (25.0–45.0 Hz), and high (45.0–100 Hz) bands. To remove electrode-pop artifacts, each derivation was normalized into z-scores. Epochs containing large EEG artifacts (absolute z-score > 30) were rejected and excluded from further analysis.

Each of the four behavioral metrics (flat spots, video BMs, definite BMs, and lapses) were decimated to a resolution of 1 Hz and provided a proxy for differentiation between the states of not-lapsing (0) and lapsing (1). The mean and standard deviation of the EEG power for a particular band during the non-lapsing samples was calculated for each of the behavioral metrics. These values were then used to transform EEG power in each band during lapses to spectral z-scores and averaged across all lapse samples to obtain a mean z-score for each band. An overall mean z-score for each band for each subject was calculated by averaging over all channels and across both sessions.

Statistical analysis

Unless otherwise stated, paired t-tests were used for statistical tests. Mixed model ANOVAS with repeated measures were used to investigate differences in lapse metrics and EEG power. The data was tested for sphericity and, if violated, significance values were adjusted by the Greenhouse–Geisser correction. The rate and duration of flat spots, video BMs, and tracking error were separately analysed with session (1 versus 2) and epoch (1–28) as within-subjects factors and subject as the random factor.

RESULTS

Based on the conservative definite BM rate estimate, eight of the 15 subjects had one or more definite BMs during their two 1-h sessions. Six subjects had 30 or more definite BMs and two subjects had 123 and 144 definite BMs over the two 1-h sessions. Table 1 gives a summary of mean, standard error, and ranges of event rate and mean duration of flat spots, video BMs, definite BMs, and lapses for all subjects, and for the eight subjects who had at least one definite BM. For subjects who had one or more definite BMs, the first occurred 22.3 ± 3.4 min into the session (mean ± SE, range 5.5–51.3).

The rates and mean durations of flat spots, video BMs, definite BMs, and lapses for individual subjects are shown in Fig. 3. This shows considerable inter-subject variability in the rate of events and, to a lesser extent, in the duration of events. Fig. 4 shows a histogram of the distribution of the duration of lapses and definite BMs. This indicates that the most frequent definite BMs are of 2–3 s duration and that the rate of longer events decreases approximately exponentially reaching a plateau of less than 1 h^{-1} for durations longer than approximately 9 s.

Over all 15 subjects, 2.0 ± 1.0% (range 0.0–12.6%) of each session was classified as definite BM and 5.0 ± 2.0% (range 0.0–26.3%) in the state of lapsing.

There was no correlation between lapse rate and age ($r = -0.012$, *t*-test, $P = 0.97$), nor between the proportion of the task spent lapsing and age ($r = -0.05$, *t*-test $P = 0.86$). This indicates that propensity to lapse is not related to age, at least within the range of ages studied. However, this does not imply that the incidence of lapsing would not be related to age over a wider age range.

There was no difference between sessions 1 and 2 in terms of the number of subjects who had a definite BM (5 versus 8, McNemar's test $P = 0.25$) and only a marginal increase in the rate of definite BMs (10.8 versus 19.6 h^{-1}, $P = 0.085$).

The rate of flat spots without a coincident video BM was 5.0 ± 1.6 events per hour, whereas the rate of video BMs without a coincident flat spot was 19.1 ± 7.1 events per hour. Visual analysis of the tracking and video data (e.g. Fig. 2) indicated that this difference is primarily because of the flat spots being a conservative lapse estimate; many of these video BMs coincided with a slowly drifting, non-coherent tracking response. There was no difference between flat spot and video BM event durations ($P = 0.540$). However, there was a marginal difference between the event rates ($P = 0.058$).

After completing the tracking, subjects were asked whether they were aware of having any 'lapses' and/or sleep episodes during the session. All but one of the 14 subjects who had at least one lapse in a session considered that they had lapsed during that session. Conversely, four subjects considered they had lapsed during a session in which no lapses were evident.

From the video rating, samples were classified as (mean ± SE): alert = 57.5 ± 8.4%, distracted = 0.2 ± 0.1%, forced eye closure = 0.3 ± 0.1%, light drowsy = 28.6 ± 6.0%, deep drowsy = 9.1 ± 3.3%, and video BM = 4.2 ± 1.7%. Flat spot rates (mean ± SE) during each video rating level were alert = 0.8 ± 0.4 h^{-1}, distracted = 130.6 ± 125.8 h^{-1}, forced eye closure = 60.5 ± 40.2 h^{-1}, light drowsy = 21.6 ± 7.0 h^{-1}, deep drowsy = 67.3 ± 20.0 h^{-1}, and video BM = 179.3 ± 27.4 h^{-1}. Repeated measures ANOVA showed a strong main effect of video-rating level on flat spot rate ($F = 21.007$, $P < 0.0001$).

Fig. 5 shows the time course of the rate and duration of flat spots and video BMs for both sessions of all subjects ($n = 15$) over 128-s epochs. As expected, a gradual increase in the flat spot and video BM rates was observed with time-on-task until the middle of the session. Conversely, the monotonic decrease in the event rates in the latter half of both sessions was unexpected. Event durations were found to be relatively stable throughout the sessions. Tracking error, defined as the absolute difference between the target and user response, and averaged over the 128-s equally difficult target epochs, is shown in Fig. 5c. This shows a similar time course to the flat spots and video BMs.

Table 1 Mean, standard error, maximum, and minimum values of event rate and duration of flat spots, video BMs, definite BMs, and lapses. Results are shown for all subjects ($n = 15$) and for those who had definite BMs ($n = 8$). As all subjects did not have flat spots or video BMs, the mean event duration was calculated over different n for each type of event

	All subjects, mean ± SE (min, max)	*Definite BM subjects, mean ± SE (min, max)*
Event rate (h^{-1})		
Flat spots	23.3 ± 8.9 (0.0, 102.5)	43.3 ± 13.2 (9.5, 102.5)
Video BMs	35.1 ± 12.2 (0.0, 142.0)	65.1 ± 16.8 (8.5, 142.0)
Definite BMs	15.2 ± 5.9 (0.0, 72.0)	28.5 ± 8.8 (2.0, 72.0)
Lapses	39.3 ± 12.9 (0.0, 141.5)	72.5 ± 16.9 (16.0, 141.5)
Event duration (s)		
Flat spots	3.4 ± 0.4 (1.9, 5.4); $n = 10$	3.8 ± 0.4 (2.1, 5.4)
Video BMs	3.4 ± 0.5 (1.0, 7.7); $n = 12$	4.0 ± 0.7 (1.9, 7.7)
Definite BMs	3.2 ± 0.4 (1.8, 4.6); $n = 8$	3.2 ± 0.4 (1.8, 4.6)
Lapses	3.4 ± 0.5 (1.0, 8.3); $n = 14$	4.4 ± 0.7 (2.0, 8.3)

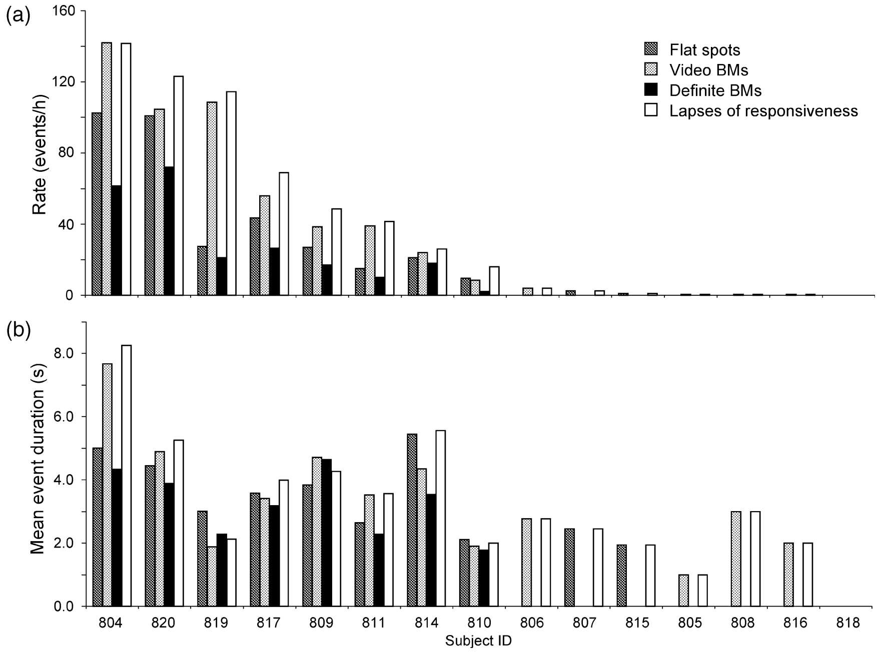

Figure 3. Inter-subject variability: (a) rates and (b) mean durations of flat spots, video BMs, definite BMs, and overall lapses during the tracking task (over two 1-h sessions). Subjects are ordered in descending order of lapse rate.

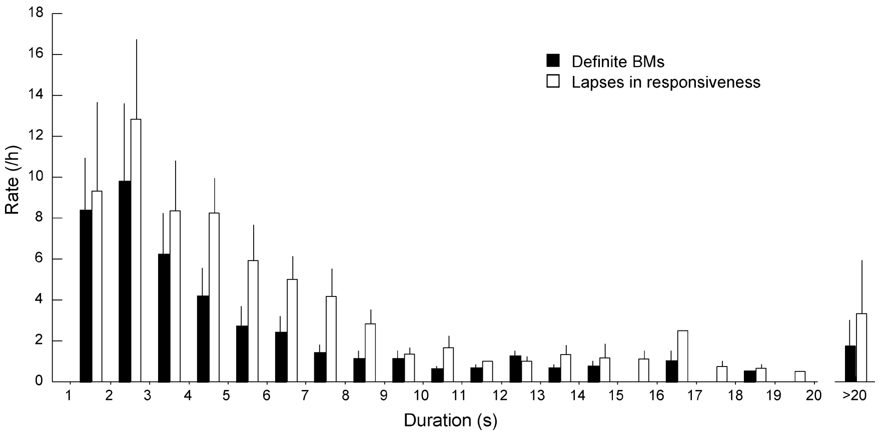

Figure 4. A histogram showing the distribution of the durations (mean ± SE) of definite BMs (filled bars) and lapses (unfilled bars). The bin width is 1.0 s. The final pair of bars depict lapses of >20 s duration.

The sphericity assumption was not met and, hence, Greenhouse–Geisser correction was applied to the mixed model ANOVA with repeated measures. A main effect of session was only observed in the flat spot rate ($F = 4.621$, $P = 0.050$). Tracking error varied with epoch (main effect; $F = 4.102$, $P = 0.022$) and showed a quadratic trend ($F = 7.550$, $P = 0.016$). Similarly, the video BM rate also varied with epoch (main effect, $F = 2.810$, $P = 0.050$) and showed a quadratic trend ($F = 8.326$, $P = 0.012$), confirming the apparent improvement in performance toward the end of the session. There was no interaction between session and epoch for any measure.

Fig. 6 shows the mean z-scores for power in each EEG spectral band during flat spots, video BMs, definite BMs, and lapses, relative to non-lapsing. It indicates a mean increase in power with lapses in the delta, theta, and alpha bands, and a decrease in the beta, gamma, and high-frequency bands. Mixed model repeated measures ANOVA, with metric (flat spots, video BMs, definite BMs, and lapses) and band (delta, theta, alpha, beta, gamma, and high) as factors, subject as random factor, and power z-score as the dependent variable showed, a main effect of band ($F = 9.847$, $P = 0.002$) but not metric ($F = 0.506$, $P = 0.581$). Within-subjects contrasts showed linear ($F = 25.455$, $P = 0.001$) and cubic ($F = 7.664$, $P = 0.028$) trends for band. There was no interaction between metric and band.

We also calculated the mean power z-score in each EEG spectral band for flat spots that did not overlap with video BMs. We observed a scaled down version of the pattern for all flat spots shown in Fig. 6, with a less positive mean z-score for delta, theta and alpha bands and a less negative mean z-score for beta, gamma, and high-frequency bands. A separate mixed model ANOVA with repeated measures showed that there was a main effect of flat spot Type (all flat spots versus flat spots without video BMs, $F = 5.116$, $P = 0.050$) and an interaction between type and band ($F = 4.991$, $P = 0.013$). There was a

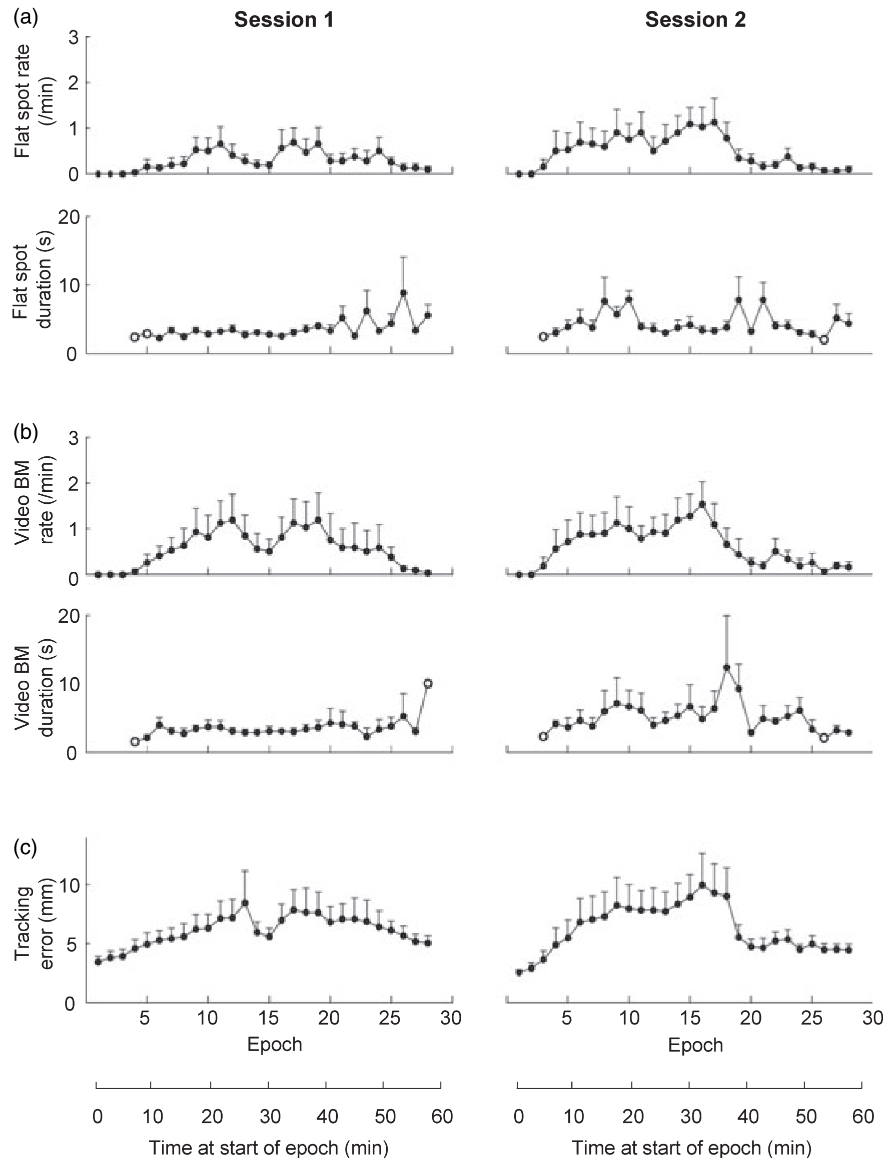

Figure 5. Time course of performance metrics (mean ± SE) used to determine lapse characteristics for all subjects for sessions 1 and 2: (a) flat spot rate and duration, (b) video BM rate and duration, and (c) tracking error. The epoch size was 128 s corresponding to a cycle of the periodic pseudo-random target. A hollow circle indicates epochs for which SE could not be calculated as only one subject contributed to the mean.

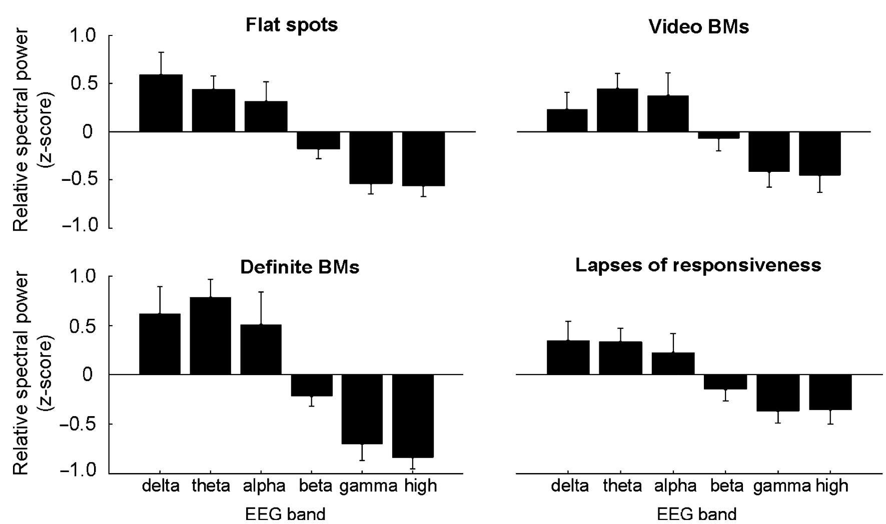

Figure 6. Changes in EEG spectral power during lapses in terms of *z*-scores (mean ± SE) relative to power during the non-lapsing state averaged across all channels, both sessions, and for all subjects during flat spots, video BMs, definite BMs, and lapses of responsiveness.

linear trend across band ($F = 9.848$, $P = 0.012$) and also an interaction in the linear trend between flat spot type and band ($F = 9.087$, $P = 0.015$), with the gradient of the trend for flat spots without video BMs being 53% smaller than that for all flat spots. This agrees with our observation that the power z-scores for flat spots were scaled down by approximately half when video BMs were not present.

EEG power showed a correlation with lapses in the delta ($r = +0.06 \pm 0.03$, $P = 0.040$), theta ($r = 0.06 \pm 0.03$, $P = 0.009$), gamma ($r = -0.07 \pm 0.02$, $P = 0.006$), and high ($r = -0.09 \pm 0.03$, $P = 0.004$) bands. Similar correlations were obtained for definite BMs: delta ($r = +0.09 \pm 0.04$; t-test, $P = 0.04$), theta ($r = +0.13 \pm 0.04$; $P = 0.007$), gamma ($r = -0.11 \pm 0.02$; $P = 0.003$), and high ($r = -0.13 \pm 0.02$; $P < 0.001$). However, the mean absolute correlation across bands between lapses and definite BMs was not different ($P = 0.224$).

DISCUSSION

This is the first study to investigate the characteristics of complete lapses of responsiveness during an extended CTT in non-sleep-deprived subjects. From this, we have been able to demonstrate frequent unequivocal lapses in a substantial proportion of normal non-sleep-deprived subjects during normal working hours. Our study has also shown increases in both lapsing and tracking error during the first 30 min of a 1-h session but subsequent decreases over the second 30 min, despite the absence of external temporal cues.

The high incidence of lapsing in young, healthy, and non-sleep-deprived adults carrying out a continuous visuomotor task during normal work hours was unexpected. While the experimental conditions were intentionally conducive to sleep (i.e. warm room, lights down, quiet environment, post-lunch, circadian low), our requirement of a continuously changing motor response, being recorded on video, being in an overt experiment, having been instructed to attend to the tracking task as accurately as possible at all times, and that subjects not be sleep-deprived, could well have prevented most, if not all, lapses. If these factors did have an inhibitory effect, it certainly was not sufficient to prevent lapsing in nearly all of our subjects.

Most previous studies have used sleep deprivation as a means of inducing and/or increasing the likelihood of a subject lapsing (Cajochen *et al*., 1999; Dinges *et al*., 1997; Doran *et al*., 2001; Gillberg and Åkerstedt, 1998; Roge *et al*., 2003). However, Van Orden *et al*. (2000) reported that nine out of 15 non-sleep-deprived participants had performance lapses in a 53-min long 2D compensatory tracking task, although the rate and characteristics of these lapses were not reported.

The rate of video BMs was considerably higher than that for definite BMs (1.97 times), which appears to be primarily because of BMs causing events other than flat spots in tracking, such as the response cursor drifting away from the target. Conversely, some flat spots occurred without coincident behavioural signs on video, which appears to reflect a combination of the conservative approach taken to rating the video (e.g. questionable lapses were rated deep drowsy) plus an absence of sleep-related signs with some lapses of performance. It is unclear whether these latter lapses are arousal-related microsleeps or non-arousal-related lapses of sustained attention or diverted attention, although we can eliminate overt-diverted attention as this was identified in our video rating.

Nearly all subjects who had one or more probable lapses were aware that they had lapsed during the session. This agrees with anecdotal evidence of drivers being aware that they have briefly 'nodded off' while driving at night. However, four subjects considered they had lapsed during one or more sessions in which no lapses were evident. This may indicate that some subjects have a tendency to over-estimate lapses but could also reflect non-arousal-related lapses in attention.

On both sessions, the mean absolute tracking error and rates of flat spots and video lapses were all seen to increase progressively up to around the middle of the session but then decrease over the second half. This improvement in performance is surprising considering there was no break in the task, and subjects were not explicitly aware of the remaining task time. However, humans have been shown to possess an ability to estimate time on a 1-h scale (Aschoff, 1985, 1998). Therefore, we speculate that subjects may have been able to gauge the approximate time remaining and to suppress or reverse the effects of drowsiness in an anticipation of the end of the test. A similar trend was observed by Van Orden *et al*. (2000) where the mean tracking error of subjects performing a 2D compensatory tracking task increased until approximately 12 min into the task, followed by a plateau or possibly a decrease until the end of the 53-min test. We have been able to show definitively that performance improved towards the end of the 1-h tracking session, despite the absence of external temporal cues. Another intriguing possibility is that lapses during a task act as mini 'rest periods' for the subject and that the cumulative effect of such lapses may have a restorative effect, leading to reduced drowsiness and improved performance.

The importance of using a substantial duration (1 h) for the tracking task is demonstrated by the onset time of the first definite BM in a session which ranged from 5.5 to 51.3 min. The mean time of 22.3 min for the occurrence of the first definite BM is comparable with that seen by Thiffault and Bergeron (2003) in a study of driver fatigue in which they found that the impact of fatigue is robust and appeared quite early during each driving session, with a marked peak occurring after 20–25 min of driving. Shorter tests of sustained vigilance, such as the PVT (Dinges and Powell, 1985), typically of 10-min duration, might fail to induce BMs in subjects who are not sleep-deprived because of the relatively short duration of time the subject is required to be attentive. Other studies have suggested that signs of fatigue may only be observed in some subjects after about 60 min of driving (Skipper and Wierwille, 1986) or on a vigilance task (Galinsky *et al*., 1993).

It needs to be emphasized that the level of lapsing seen in the current study is unlikely to be paralleled in the somewhat equivalent on-road task of driving a vehicle. Fatigue and boredom are likely to be more evident and more difficult to counter in a monotonous task such as the CTT as opposed to on-road driving. Degraded performance and level of sleepiness in such real-life situations are likely to be substantially less because of the higher level of stimulation (Åkerstedt *et al.*, 2005) and the far greater consequences of lapses. This notwithstanding, numerous studies have demonstrated a convincing link between tiredness, fatigue, and falling asleep and fatal accidents on the road (Cummings *et al.*, 2001; Horne and Reyner, 1995; Lal and Craig, 2001; Sagberg, 1999; Stutts *et al.*, 2003).

The presence of BMs was determined conservatively by a requirement for coincident and independent video-based BMs and non-response tracking. Thus, irrespective of uncertainties about underlying mechanisms of microsleeps in terms of phasic changes in arousal and/or attentional systems in the brain, their frequent occurrence in normal non-sleep-deprived subjects on sustained continuous tasks is indisputable. Although eye fatigue is likely during an extended tracking task, as are consequent relief measures such as temporary closing of eyes, we do not consider these a confounding factor in causing BMs. Our video rating scale had a category of 'forced eye closure' to take such events into consideration and to differentiate them from potential video BMs. We have also noted that forced eye closures of even a few seconds do not cause deterioration in tracking performance as subjects are able to plan and execute predictive movements to maintain the cursor on a low-bandwidth preview target.

Changes in EEG power with poorer performance and with reduced levels of alertness have been seen previously (Huang *et al.*, 2001; Jung *et al.*, 1997; Makeig and Inlow, 1993; Makeig and Jung, 1995, 1996; Makeig *et al.*, 2000; Santamaria and Chiappa, 1987). Our EEG analysis extends this by showing significant changes in the EEG spectral bands during flat spots, video BMs, definite BMs, and lapses, compared with the non-lapse state. The changes involved increases in delta, theta, and alpha power and decreases in beta (small), gamma, and higher frequency powers. Being most definitively related to the level of arousal, it is not surprising that changes were most pronounced during definite BMs. However, the correlations between EEG band power and definite BMs were, at best, low. The bands showing the greatest mean correlation with definite BMs were theta, gamma, and higher frequency activities. Theta power has been previously shown to correlate with auditory alertness (Huang *et al.*, 2001). Lal and Craig (2005) has reported slow wave activity (delta and theta) to be correlated with driver fatigue.

We observed similar trends in EEG power when flat spots occurred without apparent video sleep, although less pronounced. Hence, these are likely to represent 'shallower' lapses. This seems particularly likely given our observation that lapses typically begin with flat spots, with video BM appearing slightly later. The trend may also reflect our conservative identification of video BMs.

Despite some initial apprehension as to whether subjects would lapse at all during the 1-h tracking task, and the conservative procedure used to identify both BMs and lapses, this study has demonstrated that serious lapses can occur in young healthy non-sleep-deprived adult males to a much greater extent than previously recognized. This has major implications for occupations that require sustained alertness over long periods of time. From a safety point of view, it would clearly be very desirable to counter such lapses by early detection and wake-up systems. We are working towards the development of a lapse detection system based upon subtle indicators of lapses in the EEG (Davidson *et al.*, 2005; Peiris *et al.*, 2005).

Obvious questions arising from our study are: Is the incidence and characteristics of lapsing any different in women than men? And, although we found no age effect over the age range of 18–36 years, does the likelihood of lapsing increase, as might be expected, at older ages? These questions are being addressed in a follow-up study.

ACKNOWLEDGEMENTS

This work was supported by the New Zealand Foundation for Research Science and Technology and the University of Canterbury (Doctoral Scholarship).

REFERENCES

Åkerstedt, T., Peters, B., Anund, A. and Kecklund, G. Impaired alertness and performance driving home from the night shift: a driving simulator study. *J. Sleep Res.*, 2005, 14: 17–20.

Arnedt, J. T., Geddes, M. A. C. and MacLean, A. W. Comparative sensitivity of a simulated driving task to self-report, physiological, and other performance measures during prolonged wakefulness. *J. Psychosom. Res.*, 2005, 58: 61–71.

Aschoff, J. On the perception of time during prolonged temporal isolation. *Hum. Neurobiol.*, 1985, 4: 41–52.

Aschoff, J. Human perception of short and long time intervals: its correlation with body temperature and the duration of wake time. *J. Biol. Rhythms*, 1998, 13: 437–442.

Bills, A. G. Blocking: a new principle of mental fatigue. *Am. J. Psychol.*, 1931, 43: 230–245.

Cajochen, C., Khalsa, S. B., Wyatt, J. K., Czeisler, C. A. and Dijk, D. J. EEG and ocular correlates of circadian melatonin phase and human performance decrements during sleep loss. *Am. J. Physiol.*, 1999, 277: 640–649.

Cummings, P., Koepsell, T. D., Moffat, J. M. and Rivara, F. P. Drowsiness, counter-measures to drowsiness, and the risk of a motor vehicle crash. *Inj. Prev.*, 2001, 7: 194–199.

Davidson, P. R., Jones, R. D. and Peiris, M. T. R. Detecting behavioural microsleeps using EEG and LSTM recurrent neural networks. *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2005, 27: 5754–5757.

Delorme, A. and Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods*, 2004, 134: 9–21.

Dinges, D. F. and Powell, J. W. Microcomputer analysis of performance on a portable, simple visual RT task during sustained operation. *Behav. Res. Methods Inst. Comput.*, 1985, 17: 652–655.

Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., Aptowicz, C. and Pack, A. I. Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 hours per night. *Sleep*, 1997, 20: 267–277.

Doran, S. M., Van Dongen, H. P. and Dinges, D. F. Sustained attention performance during sleep deprivation: evidence of state instability. *Arch. Ital. Biol.*, 2001, 139: 253–267.

Galinsky, T. L., Rosa, R. R., Warm, J. S. and Dember, W. N. Psychophysiological determinants of stress in sustained attention. *Hum. Factors*, 1993, 35: 603–614.

Gillberg, M. and Åkerstedt, T. Sleep loss and performance: no 'safe' duration of a monotonous task. *Physiol. Behav.*, 1998, 64: 599–604.

Harrison, Y. and Horne, J. A. Occurrence of 'microsleeps' during daytime sleep onset in normal subjects. *Electroencephalogr. Clin. Neurophysiol.*, 1996, 98: 411–416.

Hemmeter, U., Bischof, R., Hatzinger, M., Seifritz, E. and Holsboer-Trachsler, E. Microsleep during partial sleep deprivation in depression. *Biol. Psychiat.*, 1998, 43: 829–839.

Horne, J. A. and Reyner, L. A. Sleep related vehicle accidents. *Br. Med. J.*, 1995, 310: 565–567.

Huang, R., Tsai, L. and Kuo, C. J. Selection of valid and reliable EEG features for predicting auditory and visual alertness levels. *Proc. Nat. Sci. Council*, 2001, 25: 17–25.

Jones, R. D. Measurement of sensory-motor control performance capacities. In: J. D. Bronzino (Ed.) *The Biomedical Engineering Handbook*. CRC Press, Boca Raton, FL, 2000: 1–25.

Jones, R. D., Sharman, N. B., Watson, R. W. and Muir, S. R. A PC-based battery of tests for quantitative assessment of upper-limb sensory-motor function in brain disorders. *Proc. 15th Int. Conf. IEEE Eng. Med. Biol. Soc.*, 1993, 15: 1414–1415.

Jung, T., Makeig, S., Stensmo, M. and Sejnowski, T. Estimating alertness from the EEG power spectrum. *IEEE Trans. Biomed. Eng.*, 1997, 44: 60–69.

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

Kleitman, N. *Sleep and Wakefulness*. University of Chicago Press, Chicago, 1963.

Lal, S. K. L. and Craig, A. A critical review of the psychophysiology of driver fatigue. *Biol. Psychol.*, 2001, 55: 173–194.

Lal, S. K. L. and Craig, A. Reproducibility of the spectral components of the electroencephalogram during driver fatigue. *Int. J. Psychophysiol.*, 2005, 55: 137–143.

Mackworth, J. F. *Vigilance and Habituation: A Neuropsychological Approach*. Penguin, Harmondsworth, 1969.

Makeig, S. and Inlow, M. Lapses in alertness: coherence of fluctuations in performance and EEG spectrum. *Electroencephalogr. Clin. Neurophysiol.*, 1993, 86: 23–35.

Makeig, S. and Jung, T. Changes in alertness are a principle component of variance in the EEG spectrum. *Neuroreport*, 1995, 7: 213–216.

Makeig, S. and Jung, T.-P. Tonic, phasic, and transient EEG correlates of auditory awareness in drowsiness. *Brain Res. Cogn. Brain Res.*, 1996, 4: 15–25.

Makeig, S., Jung, T.-P. and Sejnowski, T. J. Awareness during drowsiness: dynamics and electrophysiological correlates. *Can. J. Exp. Psychol.*, 2000, 54: 266–273.

Maycock, G. Sleepiness and driving: the experience of U.K. car drivers. *Accid. Anal. Prev.*, 1997, 29: 453–462.

Neale, V. L., Dingus, T. A., Klauer, S. G., Sudweeks, J. and Goodman, M. An overview of the 100-car naturalistic driving study and findings. *Proc. Int. Tech. Conf. Enhanced Safety Vehicles*, 2005, 19: 1–10.

Ogilvie, R. D. The process of falling asleep. *Sleep Med. Rev.*, 2001, 5: 247–270.

Parasuraman, R. and Davies, D. R. *Varieties of Attention*. Academic Press, Orlando, FL, 1984.

Peiris, M. T. R., Jones, R. D., Davidson, P. R., Bones, P. J. and Myall, D. J. Fractal dimension of the EEG for detection of behavioural microsleeps. *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2005, 27: 5742–5745.

Priest, B., Brichard, C., Aubert, G., Liistro, G. and Rodenstein, D. O. Microsleep during a simplified maintenance of wakefulness test. *Am. J. Resp. Crit. Care*, 2001, 163: 1619–1625.

Roge, J., Pebayle, T., El Hannachi, S. and Muzet, A. Effect of sleep deprivation and driving duration on the useful visual field in younger and older subjects during simulator driving. *Vision Res.*, 2003, 43: 1465–1472.

Sagberg, F. Road accidents caused by drivers falling asleep. *Accid. Anal. Prev.*, 1999, 31: 639–649.

Santamaria, J. and Chiappa, K. H. The EEG of drowsiness in normal adults. *J. Clin. Neurophysiol.*, 1987, 4: 327–382.

Schroeder, D. J., Touchstone, R. M., Stern, J. A., Stoliarov, N. and Thackray, R. *Maintaining Vigilance on a Simulated ATC Monitoring Task Across Repeated Sessions*. DOT/FAA/AM-94/6. US Department of Transportation – Federal Aviation Administration, Washington, D.C., 1994.

Skipper, J. M. and Wierwille, W. W. Drowsy driver detection using discriminant analysis. *Hum. Factors*, 1986, 28: 527–540.

Stutts, J. C., Wilkins, J. W., Scott Osberg, J. and Vaughn, B. V. Driver risk factors for sleep-related crashes. *Accid. Anal. Prev.*, 2003, 35: 321–331.

Thiffault, P. and Bergeron, J. Monotony of road environment and driver fatigue: a simulator study. *Accid. Anal. Prev.*, 2003, 35: 381–391.

Tirunahari, V. L., Zaidi, S. A., Sharma, R., Skurnick, J. and Ashtyani, H. Microsleep and sleepiness: a comparison of multiple sleep latency test and scoring of microsleep as a diagnostic test for excessive daytime sleepiness. *Sleep Med.*, 2003, 4: 63–67.

Torsvall, L. and Åkerstedt, T. Sleepiness on the job: continuously measured EEG changes in train drivers. *Electroencephalogr. Clin. Neurophysiol.*, 1987, 66: 502–511.

Torsvall, L. and Åkerstedt, T. Extreme sleepiness: quantification of EOG and spectral EEG parameters. *Int. J. Neurosci.*, 1988, 38: 435–441.

Valley, V. and Broughton, R. The physiological (EEG) nature of drowsiness and its relation to performance deficits in narcoleptics. *Electroencephalogr. Clin. Neurophysiol.*, 1983, 55: 243–251.

Van Orden, K. F., Jung, T.-P. and Makeig, S. Combined eye activity measures accurately estimate changes in sustained visual task performance. *Biol. Psychol.*, 2000, 52: 221–240.

Wierwille, W. W. and Ellsworth, L. A. Evaluation of driver drowsiness by trained raters. *Accid. Anal. Prev.*, 1994, 26: 571–581.