

Simulator performance, microsleep episodes, and subjective sleepiness: normative data using convergent methodologies to assess driver drowsiness

Henry J. Moller^{a,*}, Leonid Kayumov^a, Eric L. Bulmash^b, Jonathan Nhan^c, Colin M. Shapiro^a

^aHuman Performance Laboratory and Sleep Research Unit, Neuropsychiatry Program, Toronto Western Hospital, University Health Network, University of Toronto, Toronto, Ontario, Canada

^bDepartment of Psychology, Queen's University, Kingston, Ontario, Canada

^cDepartment of Human Biology, University of Toronto, Toronto, Ontario, Canada

Received 25 April 2006

Abstract

Objective: Our objective was to examine a novel standardized assessment methodology of detecting impaired driving performance due to drowsiness in a normative cohort. **Methods:** Thirty-one healthy subjects with no significant sleep, medical, and psychiatric pathology were assessed in a driving simulation paradigm. Thirty-minute simulations were repeated at two-hourly intervals (i.e., at 1000, 1200, 1400, and 1600 h). Convergent data sources included drivers' subjective ratings of sleepiness and alertness, electroencephalogram-verified microsleep (MS) episodes, and a variety of real-time driving simulator performance measures such as speed, lane tracking, reaction time (RT), and off-road events (crashes). **Results:** Significant diurnal fluctuations were noted on objective measures of RT, velocity, tracking, and MS events, indicating the highest risk of impairment in the

afternoon. By contrast, subjective ratings of sleepiness and alertness did not demonstrate significant circadian variation. The mean incidence of MS episodes and crash risk correlated highly ($r=.748$). **Conclusions:** This prospective study demonstrates the relevance of multiple convergent measures for comprehensive assessment. The divergence of subjective and objective assays of impairment implies that healthy individuals may not have full insights into neurophysiologically mediated performance deficits. These results will serve as normative comparators to patients presenting with daytime somnolence and may allow a more accurate prediction of potential crash risk than noninteractive daytime polysomnogram tests such as the mean sleep latency test or the maintenance of wakefulness test.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Driving impairment; Sleep; Microsleep; Performance; Simulation; Alertness; Vigilance

Introduction

Morbidity and mortality due to driver drowsiness are major public health issues [1–3]. While medicolegal aspects of this issue are relevant to physicians and legislators, there is a relative paucity of ecologically valid hospital-based or clinic-based diagnostic instruments to assess driving impairment due to sleepiness. This issue is reflected in the recently

published revised American Academy of Sleep Medicine parameters for the clinical use of daytime polysomnogram (PSG) [4], which note that, although tests such as the maintenance of wakefulness test (MWT) can be helpful in assessing an individual's ability to stay awake when his or her inability constitutes a public safety issue, the predictive validity of the MWT and the mean sleep latency test (MSLT) for assessing crash risk is questionable. As current clinically used methodologies of assaying fitness-to-drive vis-à-vis sleep disorders remain inadequate for informing the clinical practice of somnology clinicians, the objective of this study was to propose a standardized protocol that might have increased ecological validity and that might be useable in the context of a diagnostic sleep laboratory.

* Corresponding author. Toronto Western Hospital, University Health Network, University of Toronto, 7-Main, Room 432, 399 Bathurst Street, Toronto, Ontario, Canada M4K 3H4. Tel.: +1 416 603 5667; fax: +1 416 603 5292.

E-mail address: henry.moller@uhn.on.ca (H.J. Moller).

In recent years, while there has been an increased research interest in developing predictive models of crash risk due to drowsiness using interactive driving simulator paradigms, progress has been hampered by the lack of harmonization of methodologies. While there have been reasonably established norms and standards for the diagnostic use of the MSLT and the MWT in assessing sleep proneness [5–8], there is a need to establish improved diagnostic tests of driver drowsiness that take into account ergonomic aspects of the act of driving [9]. A further concern about conventional daytime PSG monitoring with respect to driving safety is the relative insensitivity of the MSLT and the MWT PSG scoring methodology to detect impairments more subtle than the actual occurrence of an epoch of sleep (i.e., greater than 15 s of recorded sleep) [4]. In fact, even the largest sample ($n=64$) of normative MWT values by Doghramji et al. [7], with conventional sleep latency scoring as an end point, has revealed only a modest circadian variation of impaired alertness using these criteria.

The protocol in this experiment was developed in response to concerns regarding the sensitivity and the specificity of standard daytime PSG tests vis-à-vis driving. It is reasonable to assert that a performance test that assesses an individual's ability to resist sleep in a non-interactive soporific darkened environment (i.e., the MWT) may be a better methodology to assess the risk of drowsiness-related car crashes than a test where the individual is asked to intentionally initiate sleep (i.e., the MSLT). However, as a clinically useful test to guide sleep practitioners who are concerned about fitness-to-drive issues, the MWT still appears to lack significant face validity. Since driving a vehicle is a complex task of information processing that requires a variety of cognitive and psychomotor performance abilities to be intact, including alertness, attention, multitasking, memory, coordination, and visuospatial perception abilities, many of these functions are not be assessed in an MWT or an MSLT paradigm. For the case of assessing driving impairment due to drowsiness, we argue that a standardized soporific driving task has superior ecological validity to noninteractive daytime tests, as it mimics real-world circumstances where a sleepy driver might be more prone to lose consciousness and to nod off due to a low-level of stimulation or task immersion [9].

The issue of norms and standards is crucial as researchers, clinicians, and legislators move towards more ergonomically based tests of sleepiness and fatigue that are relevant to fitness-to-drive. Historically, a wide variety of driving simulator protocols have been used to research this topic, although few have been systematically integrated into an actual clinical context (i.e., a medical clinic providing assessment and treatment of sleep disorders). Further complicating this matter is the fact that, unlike the detection of impairment arising from intoxication with alcohol, impairment secondary to fatigue and sleepiness

may just as easily represent a state (e.g., brought on by sleep deprivation and excessive time driving without rest) or a trait (e.g., an enduring sleep disorder such as sleep apnea). While there have been numerous studies documenting driving impairment in the context of sleep pathologies [10–17], few have examined PSG or electroencephalogram (EEG) measurements concurrently [10]; in fact, several studies with sleep latency testing in between driving tests have been performed in clinical populations [11,16], raising methodological concerns that these periods of nap opportunities between simulation sessions might have potentially contaminated the results of both.

This implies a need for more sensitive PSG measures of sleep intrusion that could interfere with driving. With respect to the reality of sleep-related crashes, O'Hanlon and Kelley [18] have observed that it takes an unresponsive driver only 3 s to traverse a 4-m-wide highway shoulder at a speed of 60 miles/h and an angle of departure of 3°. Episodes of psychomotor unresponsiveness secondary to sleep-related lapses of alertness have been described as microsleep (MS) episodes [10,19,20]. These events can be conceptualized as a period lasting up to a few seconds during which the PSG shifts from waking to sleep-like EEG characteristics and external stimuli are not subjectively perceived; MS episodes are thought to be associated with excessive daytime sleepiness and automatic behavior or nonresponsiveness. Analogous sensitive measures of drowsiness such as alpha attenuation [21,22] and spectral analysis [23–25] have been validated, and, recently, MS event recording has been shown to be more sensitive to sleep onset than traditional PSG testing [8,20]. Unlike MSLT or MWT criteria for sleep intrusion, which require a full epoch of sleep activity to occur, MS episodes are even briefer fragments of cortical sleep activity consisting of paroxysms of sleep-related EEG activity (usually alpha or theta) intruding into wakefulness [19].

At this point, there is no clearly established time criterion defining MS duration, although Guilleminault et al. [19] originally defined *microsleep* as a “short-lasting burst of typical stage 1 sleep as described by Rechtschaffen and Kales [25] and/or a short burst of synchronous alpha/theta activity recorded.” In designing the protocol for this study, MS episodes were defined as PSG-verified episodes as short as 3 s. This was thought to be particularly relevant in light of the forensic implications of brief lapses of consciousness vis-à-vis actual crashes occurring at highway driving speeds such as those discussed above.

In this study, we screened for the occurrence of MS episodes while subjects were engaged in a driving simulation paradigm that mimicked the repeat measurement methodology of the MSLT and the MWT. The intention was to provide a methodological analogy to these standardized tests and to allow for an analysis of circadian variation in terms of both performance and neurophysiologic assessment of MS intrusion. A simultaneous measurement of PSG and simulator performance was used to obtain

a multimodal assay of performance in the context of screening for MS-related sleep intrusion. In addition, subjective levels of sleepiness and alertness were obtained to provide a comprehensive and convergent picture of potential impairment due to driver drowsiness.

In summary, our protocol examined three different methodologies that might otherwise be used separately by clinicians to assess fitness-to-drive: simulator performance, neurophysiologic monitoring, and subjective awareness of sleepiness/alertness. It is the largest controlled study of simulated driving performance with concurrent neurophysiologic sleep monitoring published to date. Normative control groups are unavailable for most simulation studies, and certainly for the simulator used in our experiment, this makes precise comparisons to other studies difficult. Thus, this study was intended to examine a normative or “ideal” performance in a cohort of healthy subjects, which could be subsequently compared to other clinical populations of interest in the future.

Methods

Following protocol approval by the hospital ethics review board, recruitment of healthy control subjects took place via advertisements on hospital and university campus bulletin boards. Interested subjects were screened to assess study eligibility prior to a clinical screening interview on the date of the study. Upon arrival at the testing site, signed informed consent for study participation was obtained from subjects prior to proceeding. Subjects completed a brief questionnaire, including the Previous Night Sleep Inventory [26] and the Epworth Sleepiness Scale (ESS) [27], to screen for abnormalities in sleep. As well, subjects completed the Center for Epidemiological Studies—Depression Scale (CES-D) [28] to screen for subjective distress due to depression symptoms that could account for any effects on performance. Following questionnaire completion, all subjects were read a standardized list of instructions relating to the driving performance task, which involved remaining in the center of the right lane and as close as possible to the posted speed limit, which ranged between 70 and 100 km/h.

PSG electrodes were applied using standard two-lead positioning for EEG, electrooculogram (EOG), and electromyogram (EMG). Principal scoring EEG leads were occipital (O2-A1 and O1-A2), with bilateral submental EMG and right supraorbital EOG leads used to corroborate the presence of sleep-related polygraphic activity as per standard PSG procedure [25]. Subjects were instructed to undertake a 10- to 15-min driving test in the simulator to become familiarized with the performance task and to control for possible learning effects. Once familiarized with the task, subjects were tested for four 30-min sessions at 2-h intervals, occurring at 1000, 1200, 1400, and 1600h, analogous to protocols for MSLT/MWT daytime testing [4]. Subjects were encouraged to ask

questions regarding the task during the practice drive, but were discouraged from interacting with the research assistant during actual testing sessions. Prior to each driving session, subjects were asked to complete instantaneous subjective ratings of sleepiness and alertness, using 10-cm lateral visual analogue scales (VAS). PSG data were collected simultaneous to the digital collection of driving performance. Following completion of the final driving session, subjects were asked to provide brief qualitative feedback comments on their participation in the experiment.

Subjects were monitored closely by a research assistant during simulation drives, with the intent to maintain homogeneous testing conditions for all drives and to avoid extraneous distractions or interruptions that could disturb the integrity of the experiment. Subjects were not allowed to nap or to consume caffeinated beverages between sessions; they were asked to remain on site between 30-min testing sessions, but were permitted to be involved in any low-key activity they wished (e.g., reading, watching TV, etc.) within the hospital. Light exposure during testing was controlled for with equal lux values throughout the day.

The York Driving Simulator (York Computer Technologies) was used to assess driving performance. The driving simulator consists of a personal computer, a 15-in. monitor, a peripheral steering wheel, accelerator, and brake accessories. The simulator has been partially validated [13–17,29,30] as an ecologically valid research tool to measure psychomotor performance. The simulator presents a forward view from the driver’s seat of a motorway road scene, with standard lane markings and sign signals appropriate to the road environment. The four-lane route has few turns, no stops signs or traffic lights, and posted speeds ranging from 70 to 100 km/h. A driver’s seat from an actual automobile was modified and added for more realistic seating during simulation. Continuous measured variables (see below) were sampled at 10 Hz and averaged over the total course of a simulation. The driving environment used for this experiment was a monotonous highway scenario, with recurrent occurrence of “virtual wind gusts,” making it necessary for the driver to periodically engage in compensatory corrective steering maneuvers to remain on the road, rather than holding the steering wheel in a rote “dead-man” position. Thus, for this protocol, wind acted as a built-in secondary performance task to prevent fixed steering as an adaptive strategy. The timing of wind gusts occurred at regular intervals of 0.0714 Hz, but was randomized by the simulator with respect to directional force (i.e., from the left to the right, or absent).

Outcome measures

As discussed, three convergent measurement techniques were used to assess sleep-related and alertness-related variations in fitness-to-drive: driving simulator performance, neurophysiologic monitoring for MS episodes, and subjective level of sleepiness and alertness prior to driving.

Driving simulator performance

Driving simulator performance was assessed using the following variables, all expressed as a mean value over the 30-min drive:

1. *Lane tracking*—measured as the deviation (in percentile) of the center of the vehicle from the center of the right-hand lane. The extreme edge of the left lane position was demarcated as 100%, and the extreme right lane position was demarcated as 0%. Thus, the ideal road position for the given task of maintaining the vehicle in the center of the right lane was 25%.
2. *Tracking deviation*—expressed as the standard deviation of road position (percentile).
3. *Speed*
4. *Speed deviation*—calculated as the mean sum of differences (in km/h) of the speed of the vehicle from the posted speed limits.
5. *Reaction time (RT)*—measured as time to the onset of corrective steering maneuvers in response to wind gusts.
6. *Crashes*—calculated as the number of times that the simulated vehicle had off-road incidents, defined by the center of the vehicle crossing over the extreme edge of the lane or by blocking of a passing vehicle in the left lane.

Neurophysiologic monitoring (MS episodes)

The number of MS episodes, recorded by ambulatory pen-and-ink PSG (Grass Technologies) during driving simulation sessions, was the primary neurophysiologically dependent variable. The criteria for PSG-related sleep staging developed by Rechtschaffen and Kales [25] were used, with the exception of the length of sleep sequences (all sequences lasting 3 s or more as opposed to the conventional 30 s or more), which was employed to evaluate MS episodes as per Guillemineault et al. [19]. Total MS episodes for each ambulatory monitoring session, consisting of characteristic alpha, theta, or true sleep EEG activity lasting 3–30 s, were evaluated, with MS lapses greater than 15 s subclassified as major MS episodes and those less than 15 s subclassified as minor MS episodes. Two independent board-certified somnologists (H.M. and L.K.) who were blinded to subject identification and

driving performance data performed post hoc scoring of MS episodes. In cases of lack of consensus or excessive EEG artifact, episodes were not scored as MS events to avoid any bias towards overscoring.

Measured subjective scales

Immediately preceding each simulated drive, subjects were asked to complete instantaneous ratings of sleepiness and alertness, using 10-cm VAS of sleepiness (VAS-S) and alertness (VAS-A). The possible range of VAS scores was 0–10.

Statistical analysis

To analyze time-of-day effects, a repeated-measures analysis was performed. Mean scores on the primary dependent variables were measured for each individual 30-min testing period, and a composite mean score was also calculated using the mean of the four testing sessions. As we were interested in comparing differences in performance and neurophysiologic and subjective measures that were repeated four times per subject, independence of data could not be assumed. Thus, paired *t* test for the continuous variables and McNemar's test for the two categorical variables (MS episodes and crashes) were used to assess within-group differences. Pearson correlation coefficients were calculated for total crash risk and all measured variables. The data were analyzed and stored using the SPSS 11.5 program (Statistical Package for the Social Sciences; SPSS, Inc., Chicago, IL).

Results

Of 34 recruited subjects, two were excluded due to abnormal sleep duration/quality on sleep screening questionnaires. One was excluded due to technical problems relating to software malfunction during testing. Furthermore, one subject did not complete the final simulation of the day due to scheduling issues. No subjects reported simulator sickness. In total, 31 healthy individuals were assessed (21 males, 10 females; mean age, 31.3±2.1 years; mean CES-D score, 7.2±1.0; mean ESS score, 6.5±0.6). Multimodal data were obtained for four separate 30-min simulation performance periods (1000, 1200, 1400, and 1600 h), with performance and neurophysiologic and

Table 1
Absolute mean values for repeat simulations: performance variables, MS episodes, and subjective sleepiness/alertness ratings

Time of day (h)	Speed (km/h)	Speed variability (km/h)	Tracking (%)	Tracking variability (%)	RT (s)	Crash	MS episodes	Subjective sleepiness (VAS-S)	Subjective alertness (VAS-A)
1000	87.5±4.7	-1.0±4.6	29.8±5.4	8.3±2.8	0.96±.45	1.2±1.7	0.6±1.1	2.7±1.7	7.3±1.3
1200	89.5±8.0	1.2±7.9	28.3±3.9	8.3±2.8	1.05±.42	1.3±1.7	0.6±1.2	2.2±1.7	7.4±1.8
1400	89.5±4.4	1.1±4.2	28.5±3.9	8.5±2.5	1.09±.39	1.6±2.0	0.9±1.8	2.5±1.8	7.6±1.6
1600	90.2±4.6	1.6±4.4	28.9±4.3	8.3±2.6	1.04±.39	1.3±1.3	1.2±1.6	2.6±2.1	7.2±2.1
Total	89.2±4.5	0.7±4.3	28.9±4.1	8.4±2.4	1.04±.39	1.3±1.2	0.8±1.2	2.6±1.4	7.4±1.5

Mean values and standard deviations for all measured variables (see Methods for further description of units used to describe each variable). Standard deviation may appear exaggerated for discrete variables with low base rates due to a skew towards zero.

Table 2
Repeated-measures comparison of variation between repeated simulation drives

Time-of-day comparison (H)	Speed	Speed variability	Tracking	Tracking variability	RT	Crash	MS episodes	Subjective sleepiness	Subjective alertness
1000 versus 1200	.091	.085	.007*	.976	.037*	.750	.798	.153	.574
1000 versus 1400	.009*	.009*	.028*	.777	.010*	.352	.625	.693	.377
1000 versus 1600	.002*	.002*	.131	.824	.039*	.541	.009*	.880	.714
1200 versus 1400	.999	.897	.592	.767	.377	.462	.484	.343	.658
1200 versus 1600	.692	.797	.209	.869	.881	.762	.014*	.131	.290
1400 versus 1600	.489	.564	.165	.469	.202	.921	.030*	.704	.068

Paired-sample *t* test comparison for continuous variables; McNemar's test for categorical variables (MS episodes and crashes).

* Indicates significance at $P < .05$.

subjective variables summarized in Table 1. A paired-sample analysis of circadian within-group differences on repeat simulations for each variable is shown in Table 2.

The following observations were made in conjunction with this analysis.

A modest trend of increasing speeds with repeat drives was noted. While speed and speed variability tended to differ significantly on morning versus afternoon drives (i.e., comparison of 1000- vs. 1400/1600-h drives yielded $P = .009/.002$ for both speed and speed variability), this difference was not significant in differentiating morning versus noontime drives, or noontime versus afternoon drives.

In further distinguishing morning versus afternoon driving patterns, mean steering RT was significantly faster on the first drive of the day compared to all subsequent drives, with $P < .05$ observed for all comparisons ($P = .037/.010/.039$ in comparing 1000-h RT values to other drives). This circadian difference in RT was not observed at a level of statistical significance in comparing the noon and afternoon drives to one another. A bar graph of mean RT values for all simulations is shown in Fig. 1.

Tracking/position analysis also showed some modest but statistically significant differences in distinguishing the 1000-h drive from the 1200- and 1400-h drives ($P = .007$ and $.028$ respectively), with the mean road position slightly deviating towards the median strip of the road on the first drive compared to subsequent drives. Crash rates did not exhibit significant diurnal variation. Overall, crashes were relatively rare events, with boxplot calcu-

lations demonstrating a relative skewedness of $.98 \pm .43$ (S.E.), as a value of "zero" was the outcome for the majority of simulations.

On neurophysiologic monitoring, MS episodes were significantly more frequent on the final afternoon drive compared to all other testing sessions ($P = .009, .014,$ and $.030$ in comparing the 1600-h driving session to the 1000-, 1200-, and 1400-h sessions respectively), as shown in Fig. 2.

Subjective ratings of sleepiness and alertness did not show significant diurnal variation in this healthy cohort, although comparison of VAS alertness self-ratings between 1400 and 1600 h approached significance ($P = .068$).

Pearson correlation coefficients were calculated for total crash risk (i.e., mean crash rate averaged over four drives) in conjunction with mean values for all variables (i.e., MS episodes as well as performance and subjective self-rating measures; *significant correlation at the .05 level, **significant correlation at the .01 level, two-tailed analysis).

The mean total crash incidence showed the following *r* values for each variable: VAS-S ($r = .315$, ns), VAS-A ($r = -.278$, ns), speed ($r = .383^*$), speed deviation ($r = .349$, ns), tracking ($r = .209$, ns), tracking variability ($r = .538^{**}$), RT ($r = .389$), and MS episodes ($r = .748^{**}$).

Discussion

This investigation was intended to serve as a study of a normative cohort for the multimodal assessment of fitness-

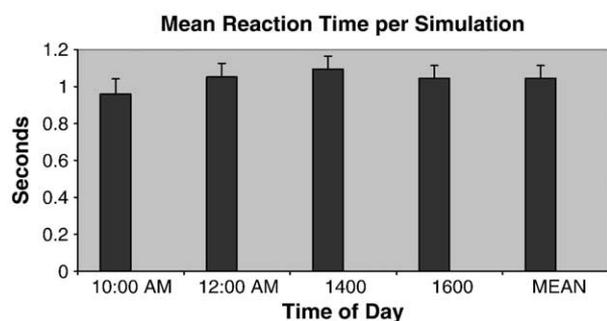


Fig. 1. Mean RT (corrective steering maneuvers, continuous digital sampling) and standard error are shown for each 30-min driving simulation and the mean of all simulations. Note the wide range used for the Y-axis.

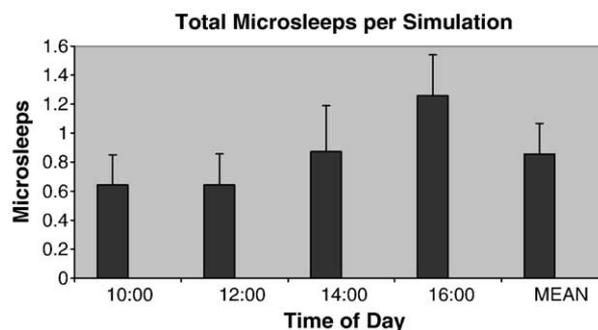


Fig. 2. Mean number of MS episodes per simulation (PSG monitoring during driving task) and standard error are shown for each 30-min driving simulation and the mean of all simulations.

to-drive vis-à-vis drowsiness. A convergent methodology, including an analysis of all three lines of data (subjective ratings, performance, and neurophysiologic data), appears as a useful methodology in obtaining a comprehensive neuroergonomic profile of the driver. While multiple testing sessions throughout the day allow for direct comparison to currently used daytime PSG tests, the advantage of our protocol is that it involves assessment of a simulated behavior approximating the “real-world” behavior of driving, rather than inferring this from noninteractive tests such as the MSLT or the MWT.

In designing a comprehensive protocol for normative standards to assess excessive sleepiness in the context of driving, it was of interest to retain some of the more potentially valuable information that might be imparted through an MWT paradigm (i.e., the objective documentation of PSG-verified sleep propensity in a prolonged soporific task, with attention to circadian variation in sleep proneness and alertness). Thus, in this study, PSG recordings of subjects were performed simultaneous to simulator performance testing as a potential correlate of impairment.

In this cohort of 31 healthy individuals, we have shown that there are clear fluctuations throughout the day in psychomotor reactivity and in the presence of EEG activity suggestive of sleep proneness. PSG monitoring during simulator testing offers a more complete and sensitive causative model of sleep-related performance changes to the clinician than simulation testing alone [11–15] or PSG recording alone using sleep-onset scoring, which requires a full epoch of sleep to occur [4]. Peak performance (lowest RT and speed variability) was noted in the morning, with a tendency towards deterioration throughout the day, and peak tendency to experience MS episodes while driving was noted in the late afternoon. It remains arguable whether performance decrement is due to true circadian variation in sleep proneness [31,32] or task fatigue (i.e., reduced ability to perform a monotonous task on repeated prolonged exposure) [33,34]. Furthermore, individual circadian differences in “morningness” versus “eveningness” have been shown [35], meaning that these general findings will not apply to all healthy test subjects.

The results in this controlled prospective laboratory protocol indicate peak crash risk in the afternoon, based on performance and MS monitoring, which would appear to correspond to the retrospective epidemiological literature analyzing time-of-day and sleep-related vehicle collisions [36–39]. Although the actual rate of collisions in this simulation paradigm did not vary with statistical significance, one must remember that crashes are rare dramatic events with a low base rate of occurrence, unlike the assessment of RT and speed variability, which consist of a continuous measurement of task vigilance. The fact that tracking variability (also sometimes referred to as lane drifting) did not show significant changes throughout the day is likely due to the recurrent nature of the secondary

performance task of corrective steering behaviors intrinsic to the RT measurement. In other studies of driver sleepiness not containing secondary tasks, tracking variability has been shown to be a possible sensitive discriminator of impairment in both simulated driving [14,16,17,40] and actual driving [41,42]. This underscores the fact that driving is a highly complex psychomotor task that should ideally be assessed multimodally, with consideration given to numerous performance and neurophysiologic variables as well as subjective experience.

In an ideal multimodal assessment, a more precise temporal association between subjective state, EEG activity signifying sleepiness, and performance testing could be established using principal components analysis [23]. A limitation of our study is that we did not analyze simultaneous temporal association between MS episodes and individual driving performance measures, as these were largely continuously collected variables. The exception to this is the variable of “crashes,” which were measured as discrete events. In future studies, it would be particularly interesting to clarify how often MS events preceded off-road events. The high positive correlation of incidence of MS episodes and crash events, higher than for any other variable pair, is striking and suggests that research exploring the probable causative role of MS episodes in contributing to crashes would be invaluable.

Also noteworthy is the relative lack of correlation of subjective self-ratings in predicting crash incidence, implying that healthy drivers may not be accurate self-assessors in terms of predicting their potential risk of collision. A limitation of this study was that subjective sleepiness and alertness ratings were obtained prior to drives. In future studies, it would be of interest to obtain numerous repeat subjective state assessments from drivers throughout the drive, although this continuous self-assessment behavior could potentially be a cognitively activating activity interfering with experimentally induced performance decrements on a monotonous task. This underscores the valuable role of continuous EEG/PSG monitoring during the driving behavior.

The finding that drivers’ own self-report ratings of alertness and sleepiness did not vary diurnally in concordance with the objective measures is intriguing, as it suggests that there may be some degree of intrinsic mismatch between objective and subjective measures of sleepiness and alertness. We propose that this is due to homeostatic regulation in healthy individuals to discount and ignore modest fluctuations in cognitive decrements. Our results indicate that, even in healthy individuals, there may be a divergence of objective and self-report measures of driver vigilance and sleepiness. This raises the issue of insight into impaired driving. Maclean et al. [1] and Arnedt et al. [13,14,40] have argued that impairment in driving performance due to sleepiness is analogous to alcohol-induced performance decrements, where, in addition to psychomotor effects on the nervous system, diminished

insight into impairment causes an additional level of risk to the driver and to the public. Alternately, lack of acknowledgment or even frank denial of unfavorable states such as disturbed wakefulness, mood, or cognition can be viewed as an adaptive response. If individuals constantly paid close attention to minor fluctuations in subjective psychophysiological states, this would be excessively distracting or even distressing. In a naturalistic setting of “real-world” driving, however, if a driver relies solely on his or her own subjective self-assessment of sleepiness and alertness, our findings suggest that he or she may not be fully aware of their decline in performance and of the growing risk of falling asleep on prolonged drives. It is important to note that these observations apply to this cohort of healthy and relatively young subjects, and it is of particular interest whether subjects with actual sleep disorders have an increased awareness of fluctuations in self-observed alertness and sleepiness, as this population is at the greatest risk for drowsiness-related collisions [1–3].

Simulator testing is, of course, only an approximation of conditions occurring in the real world. More work needs to be done in clarifying how well simulation variables such as RT or lane tracking translate into actual on-road driving situations [43]. As norms and standards of more ergonomically valid tests are developed, increasing emphasis will be put on a multimodal assessment tool using converging assessment methodologies, particularly if this could be integrated into actual commercial vehicles [44]. These normative data are intended to provide impetus for improved standards and measures in the ergonomic assessment of driver fatigue and sleepiness, as researchers attempt to bridge the gap from laboratory to roadside. Similar circadian-based simulator protocols, which allow comparison to MSLT/MWT test results, would be welcome. It clearly is informative to gather neurophysiologic and subjective data in concert with simulation data to form a comprehensive opinion of a driver’s potential impairment state due to sleepiness and fatigue, particularly if caused by underlying medical/cognitive pathology.

Ongoing research compares this cohort of normal controls to “high-risk” patients who self-identify as having excessive daytime sleepiness, using an identical protocol. It is of obvious interest how such patients perform on a soporific driving task, both with respect to MS episodes and driving performance. As well, it will be of interest to delineate the level of insight such individuals might have into any impairments. It would appear desirable for practitioners treating individuals with sleep disorders and other conditions causing neurocognitive or psychomotor impairments to have access to a standardized test with greater ecological validity than currently existing daytime tests that are recognized by legislative bodies. Finally, it is important for clinicians and researchers to communicate to legislators the limitations of current tests used to assess and to report patients at risk for driving impairment.

Acknowledgments

Funding for this research was provided to Dr. Moller by the Canadian Institute of Health Research Postdoctoral Fellowship (INMHA and Rx&D Programs).

References

- [1] Maclean A, Davies DR, Thiele K. The hazards and prevention of driving while sleepy. *Sleep Med Rev* 2003;6:507–21.
- [2] Lyznicki JM. Sleepiness, driving and motorvehicle crashes. *JAMA* 1998;279:1908–13.
- [3] Horne J, Reyner L. Vehicle accidents related to sleep: a review. *Occup Environ Med* 1999;56:289–94.
- [4] Littner MR, Kushida C, Wise M, Davila DG, Morgenthaler T, Lee-Chiong T, Hirshkowitz M, Daniel LL, Bailey D, Berry RB, Kapen S, Kramer M. Standards of practice committee of the American Academy of Sleep Medicine Practice parameters for clinical use of the multiple sleep latency test and the maintenance of wakefulness test. *Sleep* 2005;28:113–21.
- [5] Carskadon MA, Dement WC, Mitler MM, Roth T, Westbrook PR, Keenan S. Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. *Sleep* 1986;9:519–24.
- [6] Mitler MM, Gujavarty KS, Browman CP. Maintenance of wakefulness test: a polysomnographic technique for evaluation treatment efficacy in patients with excessive somnolence. *Electroencephalogr Clin Neurophysiol* 1982;6:658–61.
- [7] Doghramji K, Mitler MM, Sangal RB, Shapiro C, Taylor S, Walsleben J, Belisle C, Erman MK, Hayduk R, Hosn R, O’Malley EB, Sangal JM, Schutte SL, Youakim JM. A normative study of the maintenance of wakefulness test (MWT). *Electroencephalogr Clin Neurophysiol* 1997;103:554–62.
- [8] Banks S, Barnes M, Tarquinio N, Pierce RJ, Lack LC, McEvoy RD. The maintenance of wakefulness test in normal healthy subjects. *Sleep* 2004;27:799–802.
- [9] George CF. Driving simulators in clinical practice. *Sleep Med Rev* 2003;7:311–20.
- [10] Risser MR, Ware JC, Freeman FG. Driving simulation with EEG monitoring in normal and obstructive sleep apnea patients. *Sleep* 2000;3:393–8.
- [11] George CF, Boudreau AC, Smiley A. Effects of nasal CPAP on simulated driving performance in patients with obstructive sleep apnoea. *Thorax* 1997;52:648–53.
- [12] George CF, Boudreau AC, Smiley A. Comparison of simulated driving performance in narcolepsy and sleep apnea patients. *Sleep* 1996;19:711–7.
- [13] Arnedt JT, Wilde GJ, Munt PW, MacLean AW. Simulated driving performance following prolonged wakefulness and alcohol consumption: separate and contributions to impairment. *J Sleep Res* 2000;9:233–41.
- [14] Arnedt JT, Wilde GJ, Munt PW, MacLean AW. How do prolonged wakefulness and alcohol compare in the decrements they produce on a simulated driving task? *Accid Anal Prev* 2001;33:337–44.
- [15] Rupp T, Arnedt JT, Acebo C, Carskadon MA. Performance on a dual driving simulation and subtraction task following sleep restriction. *Percept Mot Skills* 2004;99:739–53.
- [16] Arnedt JT, Geddes MAC, MacLean AW. Comparative sensitivity of a simulated driving task to subjective, physiological, and other performance measures during prolonged wakefulness. *J Psychosom Res* 2005;58:61–71.
- [17] De Valck E, Cluydts R. Slow-release caffeine as a countermeasure to driver sleepiness induced by partial sleep deprivation. *J Sleep Res* 2001;10:203–9.

- [18] O'Hanlon J, Kelley G. Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation. In: *Vigilance*, Mackie RR, editor. New York: Plenum Press, 1977. pp. 87–110.
- [19] Guilleminault C, Billiard M, Montplaisir J, Dement W. Altered states of consciousness in disorders of daytime sleepiness. *J Neurol Sci* 1975;26:377–93.
- [20] Tirunahari VL, Zaidi SA, Sharma R, Skurnick J, 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;4:63–7.
- [21] Alloway CE, Ogilvie RD, Shapiro CM. The alpha attenuation test: assessing excessive daytime sleepiness in narcolepsy–cataplexy. *Sleep* 1997;20:258–66.
- [22] Alloway CE, Ogilvie RD, Shapiro CM. EEG spectral analysis of the sleep-onset period in narcoleptics and normal sleepers. *Sleep* 1999;22:191–203.
- [23] Makeig S, Jung TP. Changes in alertness are a principal component of variance in the EEG spectrum. *Neuroreport* 1995;7:213–6.
- [24] Lal SK, Craig A. Reproducibility of the spectral components of the electroencephalogram during driver fatigue. *Int J Psychophysiol* 2005;55:137–43.
- [25] Rechtschaffen A, Kales A. A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects. Los Angeles: University of California, Brain Information Service/Brain Research Institute, 1968.
- [26] Ladha N, Moller HJ, Kayumov L, Shapiro CM. The Previous Night Sleep Inventory (PNSI): a useful and cost-effective self-rating tool in relation to daytime somnolence and performance testing. (Abstract)20th Meeting of the Associated Professional Sleep Societies, Salt Lake City, UT, 2006.
- [27] Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep* 1991;14:540–5.
- [28] Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977; 3:385–400.
- [29] Moller HJ, Lowe A, Kayumov L, Hossain N, Shapiro C. Can impaired alertness be detected more sensitively using a computerized driving simulator? *Sleep* 2002;S252–3.
- [30] Moller HJ, Kayumov L, Bulmash EL, Shapiro CM. Assessing the effects of driver sleepiness on driving performance using combined electrophysiological monitoring and real-time computerized driving simulation. In: Vincenzi D, Mouloua M, Hancock P, editors. *Proceedings, Second International Conference on Human Performance, Situation Awareness and Automation*. Mahwah (NJ): Lawrence Erlbaum Publishers, 2004. pp. 128–33.
- [31] Buysse DJ, Monk TH, Carrier J, Begley A. Circadian patterns of sleep, sleepiness, and performance in older and younger adults. *Sleep* 2005;11:1365–76.
- [32] Pandi-Perumal SR, Seils LK, Kayumov L, Ralph MR, Lowe A, Moller H, Swaab DF. Senescence, sleep, and circadian rhythms. *Ageing Res Rev* 2002;1:559–604.
- [33] Richter S, Marsalek K, Glatz C, Gundel A. Task-dependent differences in subjective fatigue scores. *J Sleep Res* 2005;14:393–400.
- [34] Summala H, Hakkanen H, Mikkola T, Sinkkonen J. Task effects on fatigue symptoms in overnight driving. *Ergonomics* 1999;42: 798–806.
- [35] Paine SJ, Gander PH, Travier N. The epidemiology of morningness/eveningness: influence of age, gender, ethnicity, and socioeconomic factors in adults (30–49 years). *J Biol Rhythms* 2006;21:68–76.
- [36] Pack AL, Pack AM, Rodgman E, Cucchiara A, Dinges DF, Schwab CW. Characteristics of crashes attributed to the driver having fallen asleep. *Accid Anal Prev* 1995;6:769–75.
- [37] Lavie P, Wollman M, Pollack I. Frequency of sleep-related traffic accidents and hour of the day. *Sleep Res* 1986;15:275.
- [38] Kecklund G, Akerstedt T. Time of day and Swedish road accidents. *Shiftw Int Newsl* 1995;12:31.
- [39] Langlois PH, Smolensky MH, His BP, Weir FW. Temporal patterns of reported single-vehicle car and truck accidents in Texas, USA during 1980–1983. *Chronobiol Int* 1985;2:131–40.
- [40] Arnedt JT, Owens J, Crouch M, Stahl J, Carskaddon MA. Neuro-behavioral performance of residents after heavy night call vs after alcohol ingestion. *JAMA* 2005;294:1025–33.
- [41] Ramaekers JG, Robbe HW, O'Hanlon JF. Marijuana, alcohol and actual driving performance. *Hum Psychopharmacol* 2000;15: 551–8.
- [42] Verster JC, Volkerts ER, Schreuder AH, Eijken EJ, van Heuckelum JH, Veldhuijzen DS, Verbaten MN, Paty I, Darwish M, Danjou P, Patat A. Residual effects of middle-of-the-night administration of zaleplon and zolpidem on driving ability, memory functions, and psychomotor performance. *J Clin Psychopharmacol* 2002;22: 576–83.
- [43] Philip P, Sagaspe P, Taillard J, Valtat C, Moore N, Akerstedt T, Charles A, Bioulac B. Fatigue, sleepiness and performance in simulated versus real driving conditions. *Sleep* 2005;28:1511–6.
- [44] Belz SM, Robinson GS, Casali JG. Temporal separation and self-rating of alertness as indicators of driver fatigue in commercial motor vehicle operators. *Hum Factors* 2004;46:154–69.