Somnologie 12:66–74 (2008)
DOI 10.1007/s11818-008-0330-3

ORIGINALARBEIT

Murray W. Johns
Robert Chapman
Kate Crowley
Andrew Tucker

A new method for assessing the risks of drowsiness while driving

Eine neue Methode zur Einschätzung der Risiken von Schläfrigkeit am Steuer

Zusammenfassung *Fragestellung* Schläfrigkeit von Fahrzeugführern ist vermutlich eine Hauptursache von Verkehrsunfällen. Es gibt jedoch kein standardisiertes Verfahren zur Ermittlung der Schläfrigkeit zum Zeitpunkt des Unfalls. Die Schläfrigkeit wird mit einer neuen Kombination gewichteter Blinkvariablen gemessen und mit Infrarot Reflektions-Oculographie erfasst (Johns Drowsiness Skala oder JDS). Es wird untersucht, ob Leistungseinbrüche mit zwei verschiedenen Reaktionszeit-Tests und einem Fahrsimulator-Test korrelieren.

Methoden 31 gesunde Probanden absolvierten einen einfachen Reaktionszeittest (SRT) und einen Wahlreaktionszeittest (CRT) für 15 Minuten mit und ohne Schlafentzug für 27 bis 33 Stunden. Die Schläfrigkeit wurde jede Minute mit dem JDS (0–10) bestimmt. Zusätzlich fuhren 15 Probanden 70 Minuten lang mit dem Fahrsimulator mit und ohne Schlafentzug.

Ergebnisse Beim CRT und beim SRT waren die Reaktionszeiten länger und die Fehler (keine Antwort innerhalb von 2 Sekunden nach Stimulus) häufiger. Die Häufigkeit von der Strasse zu fahren nahm zu und der JDS nahm signifikant zu. Das Risiko je Minute einen Fehler im SRT zu machen und mit dem Fahrsimulator von der Strasse zu fahren nahm parallel zu einem Anstieg des JDS zu ($p < 0.001$).

Schlussfolgerung Das Risiko Leistungsfehler bei Schläfrigkeit zu machen ist mit hohen JDS Werten assoziiert. Der JDS spiegelt hauptsächlich Änderungen in der Charakteristik der Augenblinks wider. Hiermit kann die Basis für eine neue Methode gelegt werden welche die Schläfrigkeit kontinuierlich beim Fahren erfasst.

Schlüsselwörter Schläfrigkeit – Schläfrigkeit von Fahrzeugführern – Verkehrsunfälle

Summary *Question of the study* The drowsiness of drivers is believed to be a major cause of road crashes, but there is no standardized method for determining how drowsy a driver is or was at a particular time. This report describes how drowsiness, measured on a new scale (the Johns Drowsiness Scale or JDS) based on a weighted combination of ocular variables measured by infrared reflectance oculography, was related to performance failures in two different kinds of reaction-time (RT) tests as well as during simulated driving tests.

Methods 31 healthy volunteers performed simple (SRT) and choice (CRT) RT tests for 15 minutes with and without sleep deprivation for 27–33 hours. Their drowsiness was measured as a JDS score (0–10) each minute. In a separate experiment, 15 healthy young adults simulated driving in a car for about 70 minutes when alert and when sleep-deprived while their drowsiness was also measured.

Results After sleep deprivation, RTs increased and errors of omission (failure to respond within 2 seconds from the start of the stimulus) occurred more frequently in both the SRT and CRT tests, the frequency of driving off the road increased, and JDS scores also increased significantly. The

Received: 14 September 2007
Accepted: 29 January 2008

M. W. Johns · R. Chapman · K. Crowley · A. Tucker
Sleep Diagnostics Pty Ltd
Melbourne (VIC), Australia

M. W. Johns
Swinburne University of Technology
Melbourne (VIC), Australia

Dr. M. W. Johns (✉)
Suite 9, 150 Chestnut Street
Richmond, Melbourne (VIC)
3121, Australia
E-Mail: mjohns@optalert.com

K. Crowley
Melbourne University
Melbourne (VIC), Australia

risk per minute for each subject of making an error of omission in SRT tests and of driving "off road" in a car simulator increased progressively with JDS scores ($p < 0.001$).

Conclusions The risk of performance failure in the drowsy state is associated with high JDS scores, mainly reflecting changes in the characteristics of eyelid movements during blinks at the time. This could form the basis of a new method for continuously assessing the risks of drowsiness while driving.

Key words drowsiness – drowsy driving – drowsy crashes – crash risk

Potential conflict of interest statement
Sleep Diagnostics Pty Ltd is a research and development company that provided financial support for this research by way of part-time or full-time salaries for the researchers. Dr. Johns is also a share-holder of the company.

Introduction

Road crashes that are attributed to the drowsiness of drivers are often single-vehicle crashes with the driver alone in the vehicle, driving under monotonous conditions in the early hours of the morning and having been awake for a prolonged period before driving off the road and crashing [1]. Currently, there is no way of measuring a driver's drowsiness routinely while driving. Drowsiness can fluctuate rapidly, over periods of seconds, so its measurement is not as straight forward as deciding who was under the influence of alcohol at the time of a crash. The latter can be determined with reasonable accuracy by measuring the driver's blood alcohol concentration soon after the crash, but that may not be so for drowsiness.

Regulations can address some of the issues of drowsy driving, e.g. for commercial drivers, by limiting the hours of driving without a stop and the total hours of driving per day and per week [2]. People with chronic sleep disorders such as narcolepsy or severe obstructive sleep apnea, who have a high sleep propensity much of the time, can be screened and perhaps not issued with a driver's license unless their condition is treated [3]. However, circumstances evidently arise when drivers who are fit to have a license and who are alert when driving most of the time doze off at the wheel and crash, even when they obey current regulations. Those drivers are evidently not able to manage the risks of drowsy driving themselves on the basis of subjective feelings. This emphasizes the need for an objective measure of drowsiness while driving that should enable the risks of drowsy driving to be assessed in much the same way that the risks of alcohol intoxication have been [4].

We have recently described a system of infrared (IR) reflectance oculography, with IR transducers (LED emitters and a phototransistor receiver) attached to a pair of glasses that can monitor the eye and eyelid movements of drivers continuously without interfering with their performance of the driving task [5–8]. This technology has enabled several variables to be identified that can be used in combination to quantify drowsiness on a new scale, the Johns Drowsiness Scale (JDS) from 0 to 10, where 0 = very alert and 10 = very drowsy [8]. These variables include the standard deviation (SD) of the duration of blinks, the mean duration of eyelid closures (how long the eyelids remained closed and stationary), the relative velocity of eyelid closing movements during blinks, assessed by their amplitude-velocity ratios (AVRs), and the AVRs of eyelid reopening movements during blinks, all measured per minute. Other researchers have demonstrated that some of these variables, or closely related ones, change with drowsiness [9–13]. However, the AVRs for eyelid closing and reopening movements during blinks have not been used by others and provide a novel approach to the measurement of drowsiness [6].

The JDS was derived from backward stepwise multiple regression analysis, starting with many ocular variables derived from the IR reflectance oculography system as predictors of the alert and drowsy states from minute to minute. When deciding whether the subjects were indeed either alert or drowsy at particular times, it was not sufficient for them to give subjective reports of "sleepiness", for example with the Karolinska Sleepiness Scale (KSS) [14], or for us to establish that they had not slept for at least 24 hours. We used criteria based on their objective performance from minute to minute, particularly when they made errors of omission in reaction-time (RT) tests that were part of the Johns Test of Vigilance (JTV), and which also recorded their eye and eyelid movements by IR reflectance oculography during the tests.

An error of omission was defined here as the failure to make a push-button response within two seconds of the appearance of a brief change of shapes (lasting only 400 ms) on a computer screen during a 15-minute test. This definition avoids the term "lapse" as used by some other researchers [15], in which moderately delayed responses are combined with errors of omission as "lapses". We chose not to combine the many delayed or "slower responses" that drowsy subjects make (RTs between 500 and 2000 ms) in the same category as errors of omission, when there was no response.

Participants in the separate experiments that were used to develop the JDS were said to be "alert" during a JTV test if they had no errors of omission and also had < 10 % of "slower responses". For the "drowsy" condition (typically after 27–33 h without sleep), they had to have

> 5 % of errors of omission and > 15 % of "slower responses" for the whole JTV. Only those minutes of "drowsy" data that actually included at least one error of omission were used in the multiple regression. These criteria were somewhat arbitrary, but were believed to represent a level of performance impairment that would be relevant to real-life activities such as driving.

The validity of the JDS as a measure of drowsiness was later tested in other subjects who were either alert or drowsy because of sleep-deprivation. Separate investigations showed that JDS scores were significantly correlated with mean RTs in SRT tests [16], and with blood alcohol concentrations during a 6-hour period of drinking in the evening, when blood alcohol concentrations varied between zero and 0.1 % [17]. A driving simulator study involving healthy young adults, with and without sleep-deprivation by Monash University Accident Research Centre, showed that JDS scores could predict lane departure events, when all 4 wheels of the car were outside the lane within the succeeding 15 minutes, with a sensitivity of 83.3 % and specificity of 60.9 % [18].

The aim of the present study was to make a preliminary investigation of the relationship between an objective measure of drowsiness, the JDS, based on a combination of selected ocular variables, and the risks of performance failure in people who were drowsy because of sleep deprivation. Performance failure was said to occur when there was an error of omission in SRT or CRT tests, or an episode of driving with all four wheels out of the lane in the car simulator, within a particular minute. It was anticipated that this relationship might form the basis of a new method for assessing the risks of drowsiness while driving that might be applied to most if not all drivers under different circumstances, accounting for differences between subjects as well as for changes within subjects from minute to minute. Evidence was derived from two different experiments, described separately here. In the first experiment, the performance of volunteers and their JDS scores were measured each minute during two kinds of RT tests, before and after sleep deprivation. In the second experiment, the performance of other volunteers and their JDS scores were measured during simulated driving, with and without sleep deprivation.

Methods

Reaction-time experiments

Thirty one healthy volunteers (20M, 11F, mean age 26.6, range 19–33 years), whose vision was normal without correction, took part in RT tests. They were not selected or rejected on the basis of their usual sleep habits, other than them having an Epworth Sleepiness Scale score of < 11, the upper limit of normal [19]. However, participants were asked to have their usual night's sleep before tests done "without sleep deprivation". They reported the times, duration and quality of that sleep in a brief questionnaire next day. They were asked to stay up all night in the presence of another person during the period of "sleep deprivation", but their activities were not specified other than abstaining from caffeine and alcohol until after the experiment. They confirmed subjectively that they had done so before the "sleep deprivation" period. However, no objective recordings of sleep and wakefulness were made during that extended period of 27–33 hr of wakefulness to confirm their sleep deprivation. Exactly how long each participant was deprived of sleep was not an important feature of these experiments, other than as a means of increasing the likelihood of drowsiness during the RT tests and simulated driving next day.

All participants gave their informed written consent, and both experiments were conducted in accordance with the standards of the Declaration of Helsinki. The protocol for the driving simulator experiment (see below), and particularly for the sleep deprivation that it involved, was approved by the relevant Ethics Committee of Monash University, Melbourne.

Reaction-time tests

Participants performed 15-min visual RT tests of two types, simple (SRT) and choice (CRT) RT tests. In the SRT test, three circles of 20 mm diameter were displayed across a computer screen. Intermittently, they changed shape to become either squares or diamonds of similar size before reverting to circles after 400 ms. These changes occurred at random intervals between 5 and 15 s. Participants were asked to respond by pushing a button on a response pad that was held in the dominant hand as quickly as possible after any change of shape. By contrast, in the CRT test, the same stimuli were presented as in the SRT, but the participant pushed either the left or right button to indicate which of the two lateral shapes briefly became the same as the middle one. That is, the physical nature of the visual stimuli, the timing of their presentation, their duration, and the pushbutton responses were similar for both the SRT and CRT tests, but the cognitive processing required was assumed to be different according to the instructions given.

After a practice-run and familiarization with the equipment and procedures, each participant performed the SRT and CRT tests, with 30 minutes between them and in randomized order during the day after a subjectively reported normal night's sleep. They repeated those tests at the same time on the following day after missing a night's sleep, i.e. after being awake for 27–33 h. The results from the SRT and CRT tests were analyzed separately in terms of differences between subjects and the effects of "sleep deprivation". There was no interpolation of supposed RTs during errors of omission. The RT results were used mainly as evidence that the subjects did in fact have sufficient "sleep deprivation" to affect their psychomotor performance significantly. By contrast, errors of omission were interpreted as incidents of performance failure, the frequency and timing of which could be influenced by each subject's drowsiness at the time.

Drowsiness measurements

Each participant's eye and eyelid movements were monitored during the SRT and CRT tests by a system of IR reflectance oculography (Optalert™, manufactured by Sleep Diagnostics Pty Ltd, Melbourne, Australia) described elsewhere [8]. A drowsiness (JDS) score was calculated each min. This was derived from the means and SDs of several variables calculated each minute, including the duration and relative velocity of eyelid movements during blinks. These numbers were then multiplied by previously determined weightings and summed to calculate the JDS score automatically each minute during the recording, using proprietary software. Any scores >10 were made equal to 10 and those <0 were made equal to 0, so JDS scores were limited to the range 0–10. Because Optalert™ performed self-calibration during the first 3 or 4 minutes of each recording, there was an average of 11 JDS scores for each 15-minute JTV test.

The JDS scores during the SRT and CRT tests were considered separately. First, they were analyzed in terms of differences between subjects and the effects of "sleep deprivation". Then, in separate analyses, all JDS scores in "sleep-deprived" and "not sleep-deprived" conditions were combined and categorized into integer bins (0–0.9; 1.0–1.9; etc). It was anticipated that not all subjects would have JDS scores in every bin. The risk for each subject of making at least one error of omission per minute when their JDS scores were within a particular bin was calculated as follows. Considering each subject and each JDS bin separately, the number of minutes that included an error of omission was calculated as a percentage of the total number of minutes in that bin for the particular subject. The mean risk and its standard error were then calculated for all subjects with JDS scores in a particular bin, whether or not they were "sleep-deprived".

Driving simulator experiment

Fifteen volunteers (M/F = 11/4, mean age = 23.6 ± 3.2 (SD) yr, range = 21–32 yr) took part in a separate experiment performed at Monash University Accident Research Centre, Melbourne, as described elsewhere [18]. They drove in a car simulator for about 70 minutes when alert, and again after being sleep-deprived for 27–33 h, the order of tests being randomized. They were different subjects from those in another experiment using the same simulator, described in an earlier report [8]. The driver's eye and eyelid movements were monitored by IR reflectance oculography, as above. A JDS score was calculated for each minute, beginning after the first 3 or 4 minutes of driving. All JDS scores from "not-sleep-deprived" and "sleep-deprived" conditions were combined and allocated into integer bins of JDS scores, as for the RT experiments. The risk per minute was calculated for each subject of driving off the road when their JDS scores were within particular bins. This calculation was comparable to that for the risk of making an error of omission in JTVs.

Statistical methods

Separate repeated-measures ANOVAs were performed on individual RT data and JDS scores per minute, distinguishing differences between subjects and the effects of sleep deprivation. The RT data were not normally distributed and were log transformed. The distributions of JDS scores were more variable, with some approximately normal but others not, and they were not transformed. Their statistical analysis relied on the robustness of ANOVA in the presence of such differences of distribution. Spearman's ρ, Wilcoxon's τ and χ^2 tests were used where appropriate. Statistical significance was accepted when $p < 0.05$ in 2-tailed tests. Sensitivities and specificities were calculated and Receiver Operator Characteristic (ROC) Curves prepared for JDS scores as predictors, on the one hand, of errors of omission for each minute of the SRT and CRT tests, and on the other hand, for "off-road" events in the driving simulator.

Results

Reaction-time experiments

RTs recorded during SRT tests showed significant differences between subjects ($F(30, 2418) = 47.7$, $p < 0.001$) and increased RTs after "sleep deprivation" ($F(1, 2418) = 1098.8$, $p < 0.001$). The subject by "sleep deprivation" interaction was also significant ($F(30, 2418) = 19.9$, $p < 0.001$). The overall mean of RTs in the SRT tests increased from 395 ± 109 (SD) ms to 500 ± 174 ms after "sleep-deprivation". A separate ANOVA for RTs in CRT tests gave similar results, with significant differences between subjects ($F(30, 2403) = 84.3$, $p < 0.001$), increased RTs after "sleep deprivation ($F(1, 2403) = 216.2$, $p < 0.001$)

and a significant subject by "sleep deprivation" interaction ($F(1, 2403) = 15.5$, $p < 0.001$). The overall mean of RTs in the CRT tests increased from 637 ± 175 ms to 703 ± 245 ms after "sleep-deprived". Longer RTs in CRT than SRT tests was consistent with a greater cognitive work load in the CRT.

Errors of omission in SRT tests increased from 12 per 2642 stimuli (0.5 %) when "not sleep-deprived" to 151 per 2656 stimuli (5.7 %) when "sleep-deprived" ($chi^2 = 121.5$, $df = 1$, $p < 0.001$). However, only 21 subjects (67.7 %) made any such errors, whether "sleep deprived" or not. Similarly, the frequency of errors of omission in CRT tests increased from 20 per 2664 stimuli (0.8 %) to 146 per 2647 stimuli (5.5 %) ($chi^2 = 99.6$, $df = 1$, $p < 0.001$) after "sleep deprivation", those errors being made by 25 subjects (80.6 %).

In summary, this analysis of RTs provided evidence that the subjects were in fact sleep deprived to a degree that affected their psychomotor performance. Nevertheless, the majority of their responses to the visual stimuli were valid, whether or not subjects were sleep deprived. The frequency of errors of omission in SRT and CRT tests increased with "sleep deprivation", although some subjects made no such errors, either with or without "sleep deprivation". In those that did, the errors occurred intermittently during 15-minute JTVs. Thus, the risk of making an error of omission differed between subjects, was increased after "sleep deprivation", and also varied from minute to minute.

JDS scores also varied in a manner that was consistent with the effects of "sleep deprivation". JDS scores during SRT tests showed significant differences between subjects ($F(30, 302) = 91.2$, $p < 0.001$), higher JDS scores after "sleep deprivation" ($F(1, 302) = 3064.0$, $p < 0.001$), and a significant subject by "sleep deprivation" interaction ($F(30, 302) = 122.5$, $p < 0.001$). The overall mean of JDS scores during SRTs increased from 2.3 ± 2.2 (SD) to 5.0 ± 2.1 after "sleep deprivation". There were similar results from the CRT tests, with significant differences in JDS scores between subjects ($F(30,310) = 136.8$, $p < 0.001$), higher scores after "sleep deprivation" ($F(1,310) = 3764.7$, $p < 0.001$), and a significant subject by "sleep deprivation" interaction ($F(30,310) = 106.3$, $p < 0.001$). The overall mean of JDS scores during CRT tests increased from 2.3 ± 1.9 to 5.1 ± 2.4 after "sleep deprivation". Psychomotor performance, as assessed by differences in RTs and the frequency of errors of omission, was different in CRT compared with SRT tests. This was consistent with the presumed difference in cognitive work load between the two tests. However, the levels of drowsiness measured by JDS scores were very similar during both JTV types of test, performed only 30 minutes apart.

Risks for individual subjects of making errors of omission in JTVs

The risk for each subject of making at least one error of omission per minute was examined in relation to their JDS scores at the time. Data were combined from "sleep deprived" and "not sleep-deprived" sessions for all subjects, but with results from the SRT and CRT tests considered separately. The data for each minute were allocated to a JDS integer bin, between 0 and 10. The risk per minute (%) of making an error of omission was calculated separately for each subject who had JDS scores within a particular bin. The number of subjects with JDS scores in each bin varied from 0 to 23 for SRT tests. The overall mean risk and its standard error within each JDS bin is shown in Fig. 1.

There was a progressive increase in the risk of making errors of omission as JDS scores increased. This relationship involved a combination of data recorded from minute to minute in 31 different subjects, with and without "sleep deprivation". The data were not suitable for repeated measures ANOVA that would have enabled different sources of variance (e.g. subjects and conditions) to be partitioned. Instead, the overall relationship was described here in terms of Spearman's ρ which was statistically significant ($\rho = 0.43$, $p < 0.001$) with $n = 163$. This may have inflated the degrees of freedom to some extent, but the ρ would still be significant with as few as 30 degrees of freedom. The relative risk of making an error of omission each minute during SRT tests was about 10 times higher with JDS scores > 7.0 than with scores < 3.0. The comparable risks were also calculated for errors of omission during CRT tests (Fig. 2).

There was a similar trend with the CRT results to that seen with SRT test, but it did not reach statistical significance. Several subjects made errors of omission during CRT tests when ostensibly alert, i.e. without "sleep deprivation" and with JDS scores < 3.0. This serves to emphasize that not all errors of omission are caused by drowsiness, and some may be due to other factors such as distraction or stress.

Simulated driving experiment

The 15 subjects drove in the simulator for a total of 1862 minutes in two sessions, with and without "sleep deprivation". There were some missing data because two subjects did not complete their "sleep-deprived" sessions. The duration of each driving session was approximately 70 minutes, but this depended on the speed of travel twice around the circuit which simulated driving on a country road at night. JDS scores were available for a total of 1772 minutes. The overall mean of JDS scores for all subjects while driving in the simulator increased from 3.5 ± 2.1 (SD) to 5.1 ± 2.3 after sleep deprivation

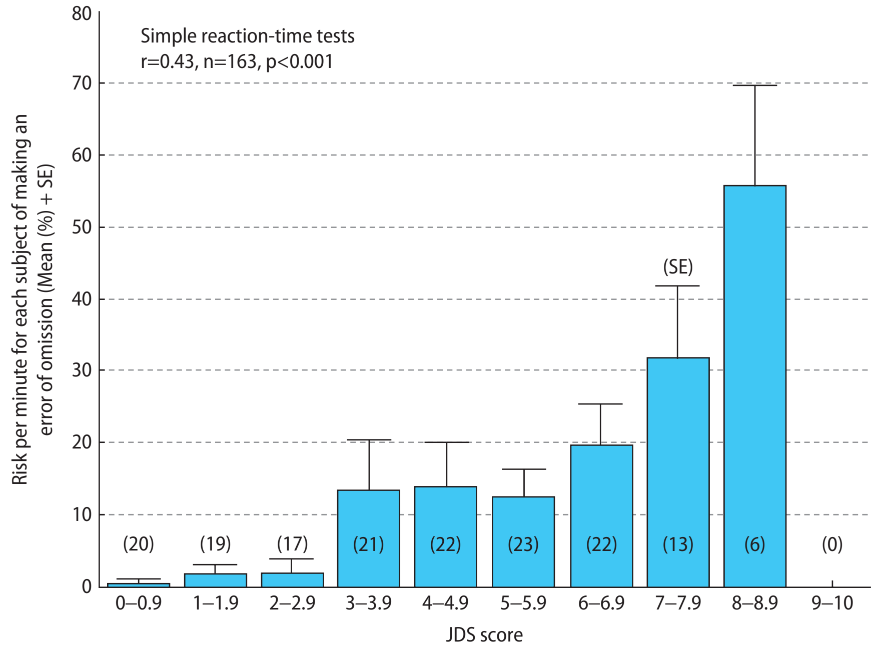


Fig. 1 Mean risk per minute (and its standard error) for each subject of making an error of omission in simple reaction time (SRT) tests vs JDS scores for "sleep-deprived" and "not-sleep-deprived" conditions combined. The number of subjects with JDS scores in each bin is shown in brackets

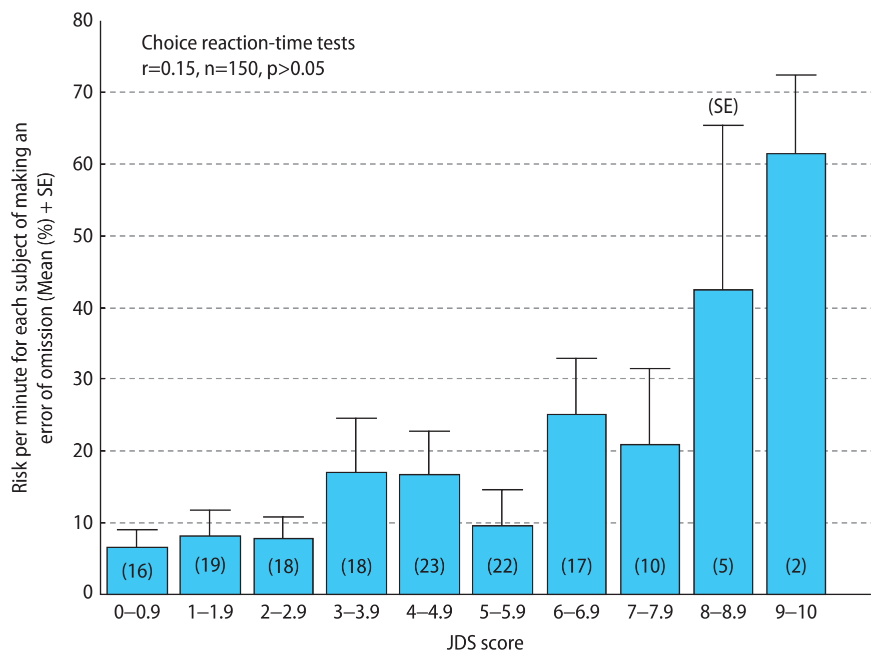


Fig. 2 Mean risk per minute (and its standard error) for each subject of making an error of omission in choice reaction-time (CRT) tests vs JDS scores for "sleep-deprived" and "not-sleep-deprived" conditions combined. The number of subjects with JDS scores in each bin is shown in brackets

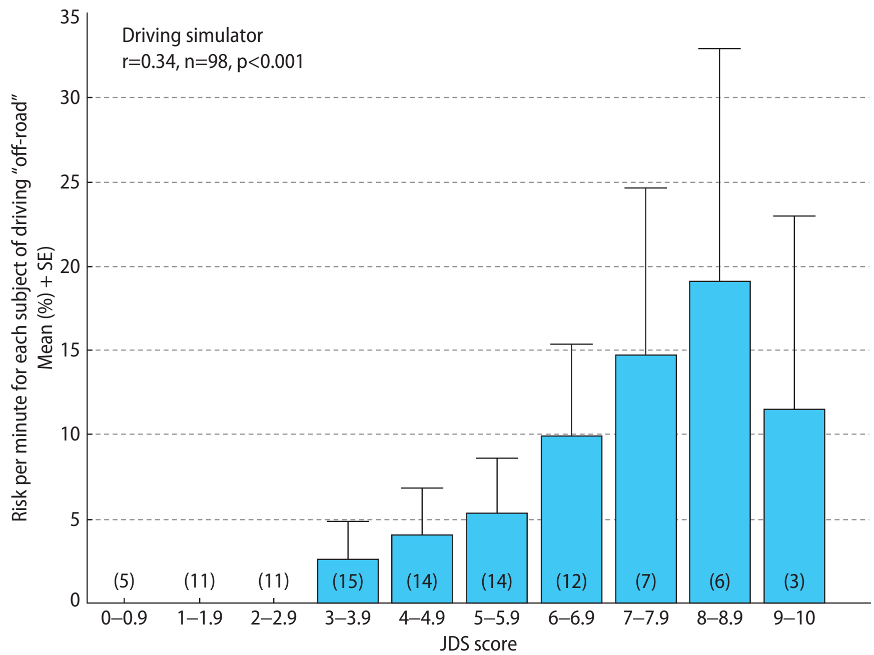
(Wilcoxon's-τ, $p < 0.001$), indicating that they were generally more drowsy while driving after "sleep-deprivation" than before it. Consistent with that drowsiness, there were many more "off-road" events with "sleep deprivation" than without it (a total of 87 vs 0, $chi^2 = 72.0$, $df = 1$, $p < 0.001$). However, only 5 (33.3%) subjects had any "off-road" events, all in the "sleep deprived" condition.

Risks for individual subjects of driving off the road

Data recorded from all subjects while driving in the simulator in both sleep conditions were combined and were allocated into JDS integer bins as was done for the JTV results. The number of subjects with JDS scores within each bin varied between 3 and 15 (Fig. 3).

The mean risk per minute for each subject of driving with all four wheels out of the lane increased as drowsiness increased, as measured by JDS scores at the time

Fig. 3 Mean risk per minute (and its standard error) for each subject of driving with all 4 wheels of the car out of the lane (i.e. "off-road") vs JDS scores for "sleep-deprived" and "not-sleep-deprived" conditions combined. The number of subjects with JDS scores in each bin is shown in brackets



(Spearman's $\rho = 0.34$, $n = 98$, $p < 0.001$). The issue about the degrees of freedom being inflated by an uncertain amount also arises here. As with the risk of making an error of omission in SRT tests, the risk of driving off the road was more than 10 times higher with JDS scores >7.0 than it was with scores <3.0. In summary, we have shown a consistent relationship between the relative risk of performance failure and JDS scores per minute in two different groups of subjects, in different test situations, with and without "sleep deprivation".

Sensitivity and specificity and receiver-operating characteristic curves for JDS scores

The sensitivity and specificity of JDS scores as predictors of errors of omission in JTV tests and of "off-road" events in the driving simulator per minute were calculated from data from all subjects in the "sleep-deprived" and "not-sleep-deprived" conditions combined. The respective ROC curves are shown in Fig. 4.

The ability of JDS scores to predict errors of omission in SRT tests was similar to that for "off-road" events in the simulator, and greater than for the prediction of errors of omission in CRT tests. The area under the respective ROC curves was 78.5 % for the SRT errors of omission, 76.1 % for the "off-road" events, and 69.7 % for CRT errors. A cut-off JDS score of 4.0 had a sensitivity of 91.0 % and specificity of 51.3 % for predicting "off-road" events during simulated driving. With a cut-off score of 7, the sensitivity was 45.0 % and the specificity 85.2 % for "off road" events. For predicting errors of omission in SRT tests, a cut-off score of 4 had a sensitivity of 89.6 %

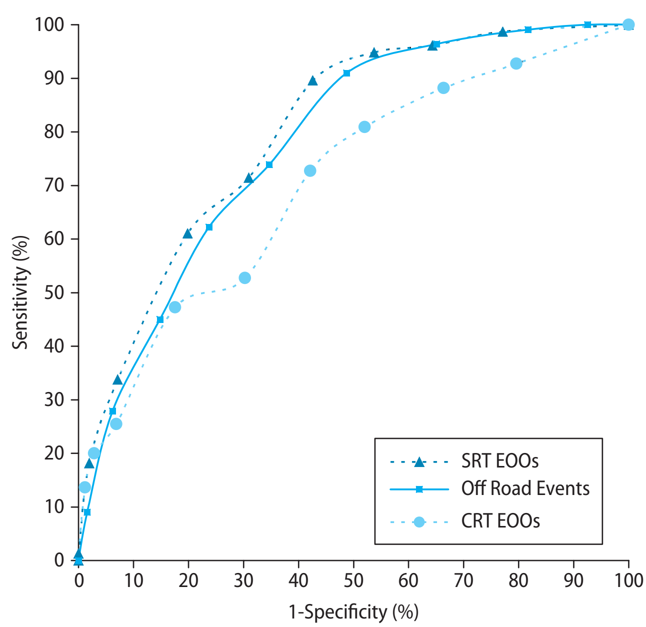


Fig. 4 Receiver-Operating Characteristic (ROC) Curves for JDS scores as predictors of errors of omission in simple reaction time (SRT) and choice reaction-time (CRT) tests, and of "off-road" events during simulated driving in "sleep-deprived" and "not sleep-deprived" conditions combined

and specificity of 57.4 %, whereas a cut-off score of 7.0 had a sensitivity of 33.8 % and specificity of 92.9 %. When choosing these cut-off scores, there is always going to be a trade-off between higher sensitivity and higher specificity.

Discussion

We have shown that, as levels of drowsiness increased during a period of "sleep-deprivation", the likelihood that there would be performance failures also increased. This was true for different groups of subjects and for different kinds of test, with errors of omission in RT tests and "off-road" events while driving in a car simulator. Subjects were generally more drowsy after missing a night's sleep (27–33 h of wakefulness), but they could still respond to visual stimuli much of the time, albeit with increased RTs. What is probably more important is that they made errors of omission (i.e. they failed to respond at all) much more frequently when drowsy. One of the important distinguishing features of a drowsy driving crash is that the driver has usually made no attempt to prevent the crash or to ameliorate its consequences, e.g. by applying the brakes or swerving to avoid hitting another object [1]. Drowsy crashes are among the most severe in terms of death, injury and property damage [1]. The implication is that drowsy drivers are simply not aware of their dangerous situation just before they crash.

In someone who is striving to remain awake, the drowsy state can fluctuate rapidly, with periods of lack of awareness of the "here-and-now" that may only last a few seconds interspersed with periods of greater alertness, more situational awareness and better performance [5]. Intermittent lack of awareness of the "here-and-now" seems to cause most errors of omission made by drowsy people in the JTV. We have previously reported on a series of 507 errors of omission in JTVs by 27 volunteers who were sleep-deprived, missing one night's sleep [20]. Their eye and eyelid movements were recorded by Optalert™ during JTVs. They could not have seen the stimulus during 46.2 % of those errors of omission because their eyes were closed at the time. However, during 26.0 % of the errors, their eyes were fully open for long enough for them to have seen the stimulus under normal circumstances. Their eyes were partially closed/open for the remaining 27.8 % of errors [20]. The drowsy state seems to be associated intermittently with periods of active inhibition of vision and of visual attention (and also of other sensory systems). That makes drowsy driving dangerous. It may also explain why the consequences of drowsy driving are to some extent a matter of chance. The consequences can be catastrophic if a period when the driver has no awareness of the "here-and-now" happens to coincide with the need for some critical response at the time. That response may be as simple as turning the steering wheel a few degrees to go round a curve in the road. However, not all episodes in drowsy subjects that involve loss of awareness of the "here-and-now" for several seconds at a time cause such performance failures.

The driving task is heavily dependent on visual input. We reasoned, therefore, that if someone's level of drowsiness was such that their risk of not responding to a clearly visible stimulus in a RT test was increased by at least ten-fold, that would represent a significant risk of critical performance failure at the time. That level of relative risk occurred in SRT tests with JDS scores > 7.0. However, the risk of performance failure began to increase initially with JDS scores > 3.0. There was a similar finding in the driving simulator, where the risk that each subject had of driving off the road per minute began to increase as JDS scores rose above 3.0, but increased rapidly above 7.0. A similar trend was seen with the CRT, but that did not reach statistical significance. These performance failures occurred after only moderate sleep deprivation (missing a night's sleep) that many people might experience, at least occasionally, for a variety of reasons.

This suggests that, even for different tasks and in different test situations, the relative risk of performance failure in different subjects increases progressively as their level of drowsiness increases, as measured by their JDS scores from minute to minute. In the context of a RT test, this means there would be high chance of not responding to what is normally a clearly visible visual stimulus. In a driving context, this may be equivalent to not seeing or responding to the presence of a curve in the road ahead, and driving off the road as a result. We suggest that an objective measure of drowsiness based on a combination of ocular variables, such as the JDS, could provide a continuous measure of the risk of critical performance failure from minute to minute while driving. When a drowsy driver's risk of performance failure (and by inference the risk of a drowsy crash) was many times higher than when he/she was alert, it could be said that the driver was no longer fit to drive safely at the time. This could form the basis of a new objective method for assessing the risks of drowsiness while driving.

Within the context of driving safety, we have assumed that it would be more important to detect the great majority of episodes of driver drowsiness than to detect all episodes when the driver was alert, i.e. it was assumed that a high sensitivity was more important than a high specificity for any particular test. Others may disagree, and a decision cannot be made on the basis of the present results alone. There is a very similar problem in deciding on a critical blood alcohol concentration (BAC) that would make a driver unfit to drive [4]. The fact that different jurisdictions have chosen quite different critical BACs (e.g. 0.05–0.08 %) highlights the difficulty in making such decisions. In terms of the measurable impairment of driving skills, a BAC of 0.05 % gives many false positive results [4].

In relation to a driver's drowsiness, we propose to circumvent this problem to some extent by using two levels of warning. The first warning would only be cautionary, suggesting that the driver pay particular attention to his/her driving and take whatever simple countermeasures were appropriate at the time [21]. On the basis of the current results, a JDS score of 4.0 might represent a level of drowsiness that would justify a cautionary warning. There would inevitably be false positive cautionary warnings at that level of drowsiness. However, a cautionary warning would not necessarily prevent the driver from continuing to drive. If drowsiness progressed to a higher level, with a much higher risk of a drowsy crash, it would trigger a critical warning suggesting that the driver should stop driving as soon as it was safe to do so. There would not be many false positive critical warnings with a cut-off JDS score of 7.0. If drivers were given a cautionary warning about their drowsiness in its early stages, they could perhaps deal with the problem before their risks of a drowsy crash became very high. However, the details of such thresholds, and decisions about making use of changes in drowsiness over the preceding few minutes, are yet to be finalized. There is no gold standard method for determining a drowsy driver's fitness to drive at a particular time against which this proposed method can be directly compared.

The present preliminary investigation has some limitations. The limited number of subjects in each experiment (31 and 15) and their age range (19–33 yr) means that the results cannot be widely extrapolated without further investigation. There were insufficient women for detailed gender comparisons to be made. In addition, it cannot be assumed that the relationship between drowsiness and "off-road" events would be the same in real-life driving as it was here with simulated driving. Nor can it be assumed that drowsiness caused by a chronic sleep disorder such as obstructive sleep apnea carries the same risk of performance failure as we have demonstrated for drowsiness caused by overnight "sleep deprivation". Further experiments are planned to address these issues. The acceptability of Optalert™ warnings to truck drivers and how they might respond to cautionary and critical warnings is currently being investigated.

Conclusions

We have measured drowsiness each minute by a method of IR reflectance oculography using a new scale, the JDS. The risk of performance failure, measured objectively in different test situations and in different groups of subjects, increased progressively as JDS scores increased. We propose that for a drowsy person to have a JDS score associated with a risk of performance failure, such as driving off the road, that was many times higher than for alert drivers, that person would be too drowsy to drive safely at the time. This could form the basis of a new objective method for continually assessing the risks of drowsiness while driving.

Acknowledgements The data from the driving simulator study were collected independently by Karen Stephan and colleagues at Monash University Accident Research Centre, Melbourne. Christopher Hocking assisted with the statistical analysis of results.

Conflict of Interest: none.

References

1. Pack AI, Pack AM, Rodgman E, Cucchiara A, Dinges DF, Schwab CW (1995) Characteristics of crashes attributed to the driver having fallen asleep. Accid Anal Prev 27:769–775
2. House of Representatives Standing Committee on Communication, Transport and the Arts. Beyond the Midnight Oil. Canberra, Australia, 2000
3. Austroads. Assessing fitness to drive. Austroads, Sydney, Australia, 2004
4. Brookhuis KA, De Waard D, Fairclough SH (2003) Criteria for driver impairment. Ergonomics 46:433–445
5. Johns MW (2003) The amplitude-velocity ratio of blinks: A new method for monitoring drowsiness. Sleep 26:A51
6. Johns MW, Tucker A (2005) The amplitude-velocity ratio of blinks: changes with drowsiness. Sleep 28:A122
7. Tucker A, Johns MW (2005) The duration of eyelid movements during blinks: changes with drowsiness. Sleep 28:A122
8. Johns MW, Tucker A, Chapman R, Crowley K, Michael NJ (2007) Monitoring eye and eyelid movements by infrared reflectance oculography to measure drowsiness in drivers. Somnologie 11:234–242
9. Morris TL, Miller JC (1996) Electrooculographic and performance indices of fatigue during simulated flight. Biol Psychol 42:343–60
10. Caffier PP, Erdmann U, Ullsperger P (2003) Experimental evaluation of eye-blink parameters as a drowsiness measure. Eur J Appl Physiol 89:319–325
11. Akerstedt T, Peters B, Anund A, Kecklund G (2005) Impaired alertness and performance driving home from the night shift: a driving simulator study. J Sleep Res 14: 17–20
12. Galley N, Schleicher R, Galley L (2004) Blink parameters as indicators of driver's sleepiness-possibilities and limitations. In: Gale A (ed) Vision in Vehicles X. Elsevier, Amsterdam, pp 189–196
13. Van Orden KF, Jung TP, Makeig S (2000) Combined eye activity measures accurately estimate changes in sustained visual task performance. Biol Psychol 52:221–240
14. Åkerstedt T, Gillberg M (1990) Subjective and objective sleepiness in the active individual. Int J Neurosci 52:29–37
15. Dinges DF, Kribbs N (1991) Performance while sleepy: Effects of experimentally-induced sleepiness. In: Monk TH (ed) Sleep, Sleepiness and Performance. John Wiley & Sons, New York, pp 98–128
16. Johns MW, Tucker AJ, Chapman R, Michael NJ, Beale CA (2006) A new scale of drowsiness based on multiple characteristics of blinks: The Johns Drowsiness Scale. Sleep 29:A365
17. Tucker AJ, Chapman RC, Crowley KE, Johns MW, Papafotiou K, Stough C (1995) The effects of blood alcohol concentration on ocular measures of drowsiness. Unpublished
18. Stephan K, Hosking S, Regan M, Verdoorn A, Young K, Haworth N (2006) The relationship between driving performance and the Johns Drowsiness Scale as measured by the Optalert system. Monash Accident Research Centre, Monash University, Melbourne, Report No 252
19. Johns M, Hocking B (1997) Daytime sleepiness and sleep habits of Australian workers. Sleep 20:844–849
20. Chapman R, Johns MW, Crowley K (2006) In the drowsy state, errors of omission in a visual reaction-time test occur with eyes open or closed. Sleep 29:A364
21. Cummings P, Koepsell TD, Moffat JM, Rivara FP (2001) Drowsiness, counter-measures to drowsiness, and the risk of a motor vehicle crash. Inj Prev 7:194–199