The Test-Retest Reliability of an Ocular Measure of Drowsiness

Johns MW [1,2], Crowley KE [2], Chapman RJ [1,2], Tucker AJ [2], Hocking CA [2]

1. Sensory Neuroscience Laboratory, Swinburne University of Technology, Melbourne, Australia
2. Sleep Diagnostics Pty Ltd, Melbourne, Australia

Introduction:

A new method has been proposed for measuring drowsiness in active people by infrared reflectance oculography (Optalert™, Sleep Diagnostics Pty Ltd, Melbourne, Australia (1,2).

This method uses a new scale, the Johns Drowsiness Scale (JDS), based on a weighted combination of variables relating to eye and eyelid movements measured every minute. A unique feature of this method is its use of amplitude-velocity ratios for measuring the relative velocity of eyelid closing and reopening movements during blinks (3-5). It also detects long eyelid closures that others have used to measure drowsiness. JDS scores vary between 0 (very alert) and 10 (very drowsy). The scale does not require calibration for individual subjects (1,2,4).

Aim:

The aim of this investigation was to measure the test-retest reliability of mean JDS scores recorded during a 15-min psychomotor vigilance test repeated within 2 hr by healthy volunteers. The tests were repeated three times under different circumstances, when alert and when drowsy because of sleep deprivation.

Methods:

The 14 healthy volunteers (M/F=10/4, ages 21-32 yr) performed a 15-min psychomotor vigilance test (the Johns Test of Vigilance, JTV) twice within about 2 hr under three different conditions. First, in a “non-sleep deprived” condition, when they performed the JTV twice on one day, at 1200 and again at 1400 hr after a normal night’s sleep reported subjectively in a questionnaire. They also performed the same JTV test after missing a night’s sleep (“sleep-deprived condition”) at 0945 hr repeated at 1130 hr, and again at 1440 hr repeated at 1545 hr. The order of “non-sleep-deprived” and “sleep deprived” days was randomised.
During the JTV test, subjects wore Optalert™ glasses to record their eye and eyelid movements, from which a measure of their drowsiness (JDS score) was derived in real-time every min.

The JTV is a laptop-based test system with a separate manual response pad. Subjects were asked to push a button as quickly as possible after they saw three circular targets on the screen change to become either squares or diamonds of the same size for only 400 msec at random intervals between 5 and 15 sec. The nature of the psychomotor vigilance test was not critical to this experiment. The JTV was used to ensure that the subjects were in the same sitting posture and engaged in the same mental and physical activity during measurements of their drowsiness. Between each pair of JTV tests subjects were also mostly sitting and taking part in the same quiet activities.

The mean of JDS scores recorded during each 15-min JTV was used as the measure of drowsiness at the time for each subject. It was the reliability of these mean JDS scores, repeated within 2 hr in the same subjects, that was being tested here in three different conditions involving different degrees of sleep deprivation, up to 34 hrs.

Results:

Fig 1 shows the overall mean of JDS scores (+ standard errors) for the 14 subjects measured twice within about 2 hr under three conditions – in the early afternoon (1200-1400 hr) after a normal night’s sleep, in the morning (0945-1130) after 27-29 hr of sleep deprivation, and again in the afternoon (1440-1545 hr) after about 33-34 hr of sleep deprivation.

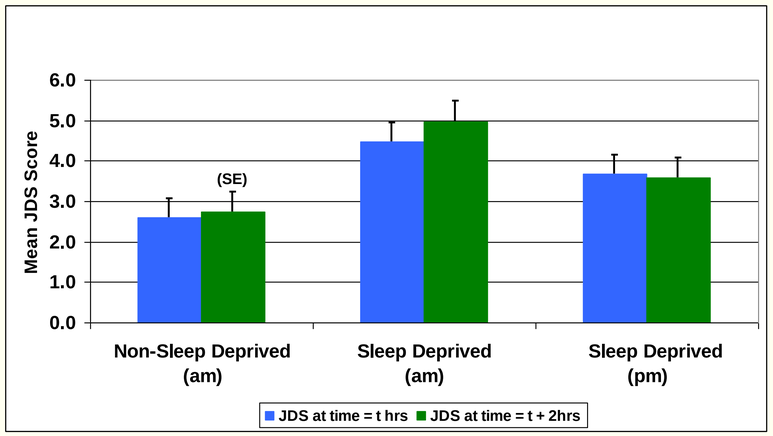

Fig. 1. Mean JDS scores in two sessions (repeated within two hours) under three different conditions of sleep deprivation.

Repeated measures ANOVA for mean JDS scores showed a significant effect for Condition (sleep deprivation) ($F(2,39) = 5.049$, $p = 0.01$), but not for Session (test-retest) ($F(1,39) = 0.980$, $p = 0.33$), or Condition by Session interaction ($F(2,39) = 0.863$, $p = 0.43$). JDS scores tended to increase with sleep deprivation and also showed evidence of a time-of-day effect, as described previously (3).

Fig 2 shows the scatterplot and regression between the mean JDS scores measured in each subject during the first and second JTV sessions in all three conditions.

A paired t-test between test and retest mean JDS scores, combining all three conditions, showed no significant difference (mean diff = 0.19 +/- 0.19 standard error, $n = 42$, $p > 0.3$), and there was a high intraclass correlation ($r = 0.80$, $n = 42$, $p < 0.001$).

References

1. M.W. Johns, A.J. Tucker, R.J. Chapman, K.E. Crowley, N.J. Michael. Monitoring Eye and Eyelid Movements by Infrared Reflectance Oculography to Measure Drowsiness in Drivers. *Somnologie* 2007; 11 (4): 234-242
2. M.W. Johns, R Chapman, K Crowley, A Tucker. A new Method for Determining When a Driver is too Drowsy to Drive. *Somnologie*, 2008; 12 (1): 66-74
3. M.W. Johns, A.J. Tucker, R.J. Chapman, N.J. Michael, C.A. Beale. A New Scale of Drowsiness Based on Multiple Characteristics of Blinks: The Johns Drowsiness Scale. *Sleep*, 2006; 29 (Suppl): A365
4. M.W. Johns, A.J. Tucker. The Amplitude-Velocity Ratios for Eyelid Movements During Blinks: Changes with Drowsiness. *Sleep*, 2005; 28 (Suppl): A122
5. A.J. Tucker, M.W. Johns. The Duration of Eyelid Movements During Blinks: Changes with Drowsiness. *Sleep*, 2005; 28 (Suppl): A122
6. R.B. Sangal, L. Thomas, M.M. Mitler. Maintenance of Wakefulness Test and Multiple Sleep Latency Test: Measurement of Different Abilities in Patients with Sleep Disorders. *Chest* 1992; 101: 898-902
7. M.W. Johns. Sleep Propensity Varies with Behaviour and the Situation in which it is Measured: The Concept of Somnificity. *Journal of Sleep Research,* 2002; 11: 61-67

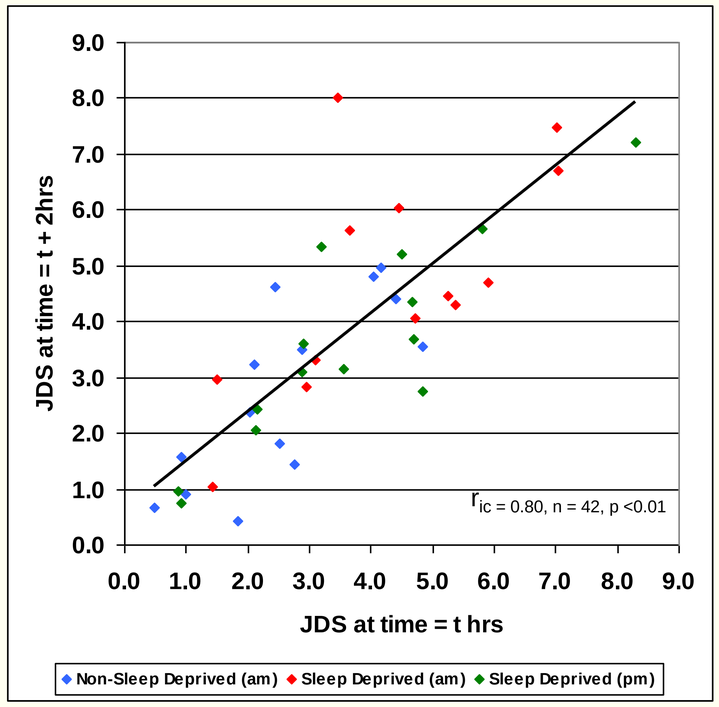

Fig. 2. Correlation between mean JDS scores in each subject, repeated within two hours, in three different conditions of sleep deprivation.

Discussion:

The measure of drowsiness provided by a mean JDS score recorded during a 15-min period is reliable in a test-retest sense when tested twice within 2 hr under similar circumstances.

The level of test-retest reliability demonstrated here for mean JDS scores ($r = 0.80$, $n = 42$, $P<0.01$) was higher than for the sleep latencies measured in consecutive naps 2-hr apart within the same subject during both the MSLT and MWT ($r = 0.61$ $n = 258$, $p < 0.001$)(6).

The mean JDS score recorded during a 15-min JTV is a reliable measure of the subject’s drowsiness at the time under those specific circumstances. It may be extrapolated to become a measure of that subject’s general level of daytime sleepiness in daily life only within the limitations inherent in all such measurements that are situation and activity-specific (7).

Conclusions:

- **The mean JDS score was reliable, in a test-retest sense, when repeated within a 2 hr period.**
- **The reliability of mean JDS scores was demonstrated here at three different levels of drowsiness induced by sleep deprivation.**
- **The mean JDS score recorded during a 15-min JTV test increased after sleep deprivation and may show a time-of-day effect.**