Real-Time Driver Drowsiness Feedback Improves Driver Alertness and Self-reported Driving Performance

Eugene Aidman¹, Carolyn Chadunow¹, Kayla Johnson¹, John Reece²
1 Defence Science and Technology Organisation, Edinburgh, South Australia, Australia
2 RMIT University, Melbourne, Victoria, Australia

1. Abstract

Background: Driver drowsiness has been implicated as a major causal factor in road accidents. Tools that allow remote monitoring and management of driver fatigue are used in the mining and road transport industries. Increasing drivers’ own awareness of their drowsiness levels using such tools may also reduce risk of accidents.

Aims: The study examined the effects of real-time oculography-generated feedback (Optalert Alertness Monitoring System - OAMS) on driver performance and levels of alertness in a Military setting.

Method: A sample of 15 Army Reserve personnel (1 female) aged 21-59 (M = 41.3, SD = 11.1) volunteered to being monitored by OAMS while they performed their regular driving tasks, including on-duty tasks and commuting to and from duty, for a continuous period of 4 to 8 weeks. For approximately half that period, OAMS-derived driver drowsiness scores were fed back to the driver in a counterbalanced repeated-measures design, resulting in a total of 419 driving periods under “feedback” and 385 periods under “no-feedback” condition.

Results: Overall, the provision of real-time OAMS feedback resulted in reduced drowsiness and improved driving performance ratings. The effect was small but remained robust after controlling for time of day and driving task duration. Both the number of JDS peaks counted for each driving session and their duration declined in the presence of OAMS feedback, indicating a pattern of drowsiness dynamics that is consistent with a genuine feedback mechanism (as distinct from random re-alerting) behind the OAMS effect.

Conclusions: Driver drowsiness can be reduced by the provision of OAMS feedback. The effect is relatively small, and varies across the 24-hour circadian cycle. Its mechanisms and practical utility have yet to be fully explored.

2. Background

Driver fatigue is a major causal factor in road accidents (1). Fatigue is also a construct that links factors such as time of day, time since waking, task duration and monotony, with safety-related outcomes (2, 3). Fatigue can result from sleepiness (drowsiness), boredom, and mental or physical exhaustion. From these causal factors, drowsiness is considered the most relevant aspect of fatigue when applied in the driving context. Driver drowsiness has been implicated in road accidents both within professional (4) and general driving
Accidents caused by driver drowsiness often occur have a similar fatality rate to alcohol-related crashes (6).

Multiple factors contribute to drowsiness, such as long working hours (7), lack of sleep (8), and medical conditions (9). Lack of sleep is more prevalent in some populations, including junior doctors (10), submariners at sea (11) and ‘fly-in fly-out’ mining workers (12). Chronic sleep restriction is a known risk factor in driving (3). It is also well established that the 24 hour circadian rhythm is marked with peaks and troughs in alertness levels as evidence by studies incorporating both subjective and objective sleepiness measures (13, 14). Task-related factors also contribute to driver drowsiness (3). These factors may include driving duration (15) and monotony (16), such as that experienced in highway driving (17).

The effects of drowsiness manifest in a reduced capacity to maintain vigilance (18). In the driving context this leads to observable changes in driver performance such as reduced capacity to maintain speed, distance between vehicles and lane-keeping (19), all of which increase the risk of road accidents (20).

With mounting evidence linking driver drowsiness to road accident risk, industry has responded with investment in driver monitoring tools aimed at mitigating this risk (21, 22). These tools employ a range of methods including those based on (a) continuous driving time, (b) specific driver performance (e.g., steering) or (c) physiological response (e.g., eye metrics). Among the latter, eye and eyelid characteristics have been used to infer levels of drowsiness (23). One of these tools, the Optalert Alertness Monitoring System (OAMS) (24) utilises infra-red (IR) reflectance oculography to monitor eyelid movements. The system uses an IR emitter and sensor mounted on a spectacle frame to continuously measure eye blink velocity, from which levels of drowsiness are derived. The OAMS has been used for detection and monitoring of driver drowsiness in the mining (25) and road transport industries (3, 26), as well as for pilot drowsiness detection in aviation (27).

OAMS has been used in commercial settings predominantly through the provision of drowsiness data to a central monitoring area. When a driver’s drowsiness reaches a predetermined risk level, interventions (e.g. taking a mandatory break) can be implemented. Applications of this type come with a significant overhead in terms of monitoring and implementation cost. In addition, such interventions rely on fatigue detection rather than prevention. The current study explores the utility of real-time drowsiness feedback to drivers for fatigue prevention. Tucker (28) has shown preliminary data indicating that providing OAMS feedback to the driver (via a dash-mounted display and auditory warnings) may reduce drivers’ drowsiness, compared to a no-feedback condition. Developing an understanding of the utility of real-time feedback to drivers may lead to a low cost strategy to improve driving outcomes in at-risk groups.

Our study aims to examine the effectiveness of OAMS in improving driver functional state. It focuses on Australian Army reserve personnel. For the majority of participants, an Army training weekend is normally preceded by a full-time working week and often a lengthy drive to the location of Army Reserve duty. In addition, members often report for duty on a weekday night, again involving the commute to and from the duty location. As a group Army
Reservists are likely to work longer hours and have less sleep when on duty (29), potentially putting them at higher risk for drowsiness related driving accidents.

We hypothesised that provision of OAMS feedback would reduce both objective drowsiness and subjective sleepiness when compared to no-feedback condition. Feedback was also expected to improve driving performance ratings.

3. Method

3.1 Participants

Fifteen Army Reserve personnel (1 female) aged from 21 to 59 years (M = 41.3, SD = 11.1) volunteered in return for partial duty time credit. Their rank ranged from Private (Sapper) to Lieutenant Colonel. The participants were members of an Army Reserve regiment who self-selected as frequent drivers from a larger study sample (30). This study was approved by the DSTO Human Sciences Ethics Committee prior to recruiting participants.

3.2 Design

The study used a repeated-measures design to test the effect of feedback (On, Off) on Johns Drowsiness Scale scores, Karolinska Sleepiness Scale scores and self-rated driving performance (lane-keeping, headway and responsiveness). The order of conditions was counterbalanced, with half of the subjects randomly selected to begin the trial with feedback on, and the remainder with feedback off.

3.3 Materials

3.3.1 Johns Drowsiness Scale (JDS)

OAMS (24) continuously monitors eye and eyelid movement and generates a drowsiness score at one minute intervals. The scores form the Johns Drowsiness Scale (JDS)(18). The JDS scale has demonstrated good test-retest reliability (31).

The OAMS comprises individually calibrated glasses, a dashboard indicator which displays JDS scores and a data collection processor. JDS scores range from 0 to 10, with higher scores indicating increasing drowsiness. A JDS score between 0 and 4.4 (inclusive) indicates a low risk level of drowsiness. A score between 4.5 and 4.9 (inclusive) indicates medium risk and JDS scores of 5.0 and above indicate a high risk (24). In the feedback condition, incursion of a JDS score into a medium or high risk range produced both auditory and visual warnings through the dashboard indicator. When a driver remained in the low risk range for 5 minutes, the indicator display blanked out but was able to be turned on by touching the screen. In the feedback-off condition, no auditory or visual warnings were available.
3.3.2 Karolinska Sleepiness Scale (KSS)

Driver alertness was measure using a 10-point modified version of the Karolinska Sleepiness Scale (32, 33). The KSS is typically used as an instantaneous measure of sleepiness (33). The version utilised in this study had participants record their current level of alertness on a 10-point scale, which is a reverse of the Kaida et al. (2007) instrument, ranging from 1 (“extremely sleepy, fall asleep all the time”) to 10 (“extremely alert”). Drivers were asked to record the start and end time of a drive as well as their alertness level using the KSS.

3.3.3 Self-rated driving performance

Participants rated their driving performance on three 5-point Likert scales: lane keeping, headway and responsiveness to road events.

3.4 Procedure

Participants were individually fitted with a pair of OAMS glasses and given a Driving Alertness Journal, which incorporated the KSS, start and end times for each drive, as well as driving performance rating scales.

Participants were monitored with OAMS while driving, for a continuous period of 4 to 8 weeks. For approximately half that period, JDS scores were fed back to the driver via a dashboard display that also generated warning sounds when the scores reached medium- and high-risk drowsiness levels. The remainder of driving tasks were performed with no feedback, while OAMS monitoring continued throughout. At the conclusion of the study, a total of 421 driving periods under “feedback” and 385 periods under “no-feedback” condition had been collected. The duration of these drive periods ranged from 1 to 180 mins, with a mean drive length of just over half an hour (26.4 minutes; SD=24.6). A linear mixed model analysis found no differences between the two conditions in drive duration, $F(1, 622.85) = 1.48, p = .46, \eta^2 < .01$, or time of day the drive was taken, $F(1, 650.44) = 1.18, p = .51, \eta^2 < .01$.

JDS scores were compared between two conditions (OAMS Feedback: On, Off), while controlling for time of day (because of natural alertness variance due to circadian rhythm) and drive duration (28, 34). Repeated measured data were analysed with mixed linear model analyses. Significance level of $p = .01$ was chosen to account for multiple comparisons.

4. Results

4.1 Effects of real-time OAMS feedback on aggregated driver state and performance

The circadian effects on JDS-measured drowsiness can be seen in Figure 1, with peaks in drowsiness during the early hours of the morning and gradually increasing throughout the day. Driver risk appears to be lowest during the day (between the hours of 7am – 5pm).
A number of analyses were conducted that examined differences between feedback conditions, using univariate linear mixed model analyses. Given the repeated measures nature of the design used, and more importantly the unequal number of repeated driving periods each participant undertook, mixed models analysis was deemed to be the most appropriate analytical approach. Participants completed each feedback condition over multiple trials, therefore both feedback condition and trial were entered as repeated fixed factors in the models. All analyses modelled the covariance structure using compound symmetry, which is the recommended approach for this design, and that demonstrated good model fit using the Akaike Information Criterion (35).

As task duration is known to influence the operator’s drowsiness levels, the initial analysis included drive length (in minutes) as a covariate. There was a significant effect of feedback on mean JDS scores, $F(1, 623.04) = 23.13, p < .001, \eta^2 = .04, 95\% CI[.01, .07]$, indicating that average drowsiness per drive was lower in the ‘feedback on’ condition ($M = 1.07, SE = 0.18)) compared with ‘feedback off’ ($M = 1.33, SE = 0.18$). The magnitude of this effect varied over the 24-hour circadian cycle. During the daytime hours (when the majority of driving occurred) the effect peaked at around 11am and declined to almost no effect between 6 and 7pm (see Figure 1). Between the hours of 10pm and 4am there were insufficient driving periods to demonstrate a clear effect of feedback.

The observed decline in the mean JDS scores in the feedback condition, was largely accounted for by the second covariate – the circadian phase represent by our time of day variable – and was no longer significant with both covariates in the model. This two-covariate model was applied to all subsequent analyses. First, it showed that the most immediate indicator of risk -
the maximum JDS score per driving session - was significantly lower in the ‘feedback on’ condition ($M = 1.97, SE = 0.21$), compared to the ‘feedback off’ condition ($M = 2.20, SE = 0.22$). The effect was small but significant: $F(1, 624.01) = 9.81, p = .002, \eta^2 = .02, 95\% CI[< .01, .04]$. We also found the subjective appraisals of safe distance keeping to be significantly higher under the feedback condition, $F(1, 270.95) = 5.34, p = .02, \eta^2 = .02, 95\% CI[< .01, .06]$. However, the effects of feedback on other appraisals of driving performance (lane-keeping and responsiveness) were not significant (see Table 1). The effect of feedback condition on KSS ratings was significant as well (Table 1), with participants rating themselves as more alert with feedback on, $F(1, 285.51) = 4.15, p = .04, \eta^2 = .01, 95\% CI[< .01, .05]$. 

### 4.2 Effects of real-time OAMS feedback on dynamic JDS metrics

Preliminary analyses, utilising aggregate characteristics of driving periods, showed that the presence of feedback had a significant impact on drivers’ subjective ratings of their safe distance keeping and sleepiness. The presence of feedback also reduced the maximum JDS scores the drivers’ reached in each driving session. In order to begin to understand the mechanism behind this difference we further examined the differences between feedback conditions, by analysing the dynamics of JDS scores over time during individual drive periods.

A significant feedback effect was found on the following dynamic JDS metrics: Number of JDS peaks, $F(1, 592.98) = 13.28, p < .001, \eta^2 = .02, 95\% CI[< .01, .05]$; JDS peak duration, $F(1, 555.59) = 4.77, p = .03, \eta^2 = .01, 95\% CI[< .01, .03]$ (See Table 1). These results indicate that the presence of feedback resulted in both reducing the number of peak JDS scores reached across a drive and decreasing the JDS peak duration.
Table 1.  
*Covariate adjusted means for outcome variables across two feedback conditions.*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Feedback Condition</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>On</td>
<td>M (SD)</td>
<td>Off</td>
</tr>
<tr>
<td>Lane Keeping</td>
<td>4.43 (0.17)</td>
<td>4.45 (0.17)</td>
<td></td>
</tr>
<tr>
<td>Safe Distance Keeping</td>
<td>4.47 (0.18)</td>
<td>4.60 (0.18)</td>
<td></td>
</tr>
<tr>
<td>Responsiveness</td>
<td>4.44 (0.18)</td>
<td>4.45 (0.18)</td>
<td></td>
</tr>
<tr>
<td>Mean JDS</td>
<td>1.18 (0.19)</td>
<td>1.34 (0.19)</td>
<td></td>
</tr>
<tr>
<td>Maximum JDS</td>
<td>1.97 (0.21)</td>
<td>2.20 (0.22)</td>
<td></td>
</tr>
<tr>
<td>KSS</td>
<td>7.59 (0.32)</td>
<td>7.30 (0.32)</td>
<td></td>
</tr>
<tr>
<td>Number of JDS Peaks</td>
<td>9.50 (0.66)</td>
<td>10.49 (0.67)</td>
<td></td>
</tr>
<tr>
<td>JDS Peak Duration</td>
<td>12.43 (0.41)</td>
<td>13.01 (0.43)</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* All descriptive values have been adjusted for the presence of driving duration and time of day as covariates.

### 5. Discussion

Our study set out to examine the effectiveness of OAMS in improving driver functional state in the context of a field trial. Overall, our results show a significant effect of OAMS feedback: its presence was consistently associated with reduced peak drowsiness (lower maximum JDS scores), improved self-reported alertness (reduced sleepiness in KSS) and drivers’ own ratings of headway. The effects are small but remain significant after controlling for known confounds, such as circadian phase (time of day as a proxy) and time on task (time since the start of each drive).

Our finding is consistent with Optalert’s own trial results (28) showing that the introduction of feedback and warnings resulted in 28% fewer medium risk warnings and 41% fewer high risk warnings when compared against the baseline, no-feedback condition.

The results confirm the predicted effects of OAMS feedback on reducing drowsiness (JDS peak scores), self-reported alertness (KSS), and improving driving performance appraisal (safe distance ratings). The magnitude of this effect varied over the 24-hour circadian cycle,
predictably peaking between 10pm and 4 am when alertness levels are at their lowest in the circadian cycle (14). The same effect declined to almost zero in between 6 and 7pm – a time window when alertness levels are typically high (13).

The important question of why the feedback condition produced these effects, remains open. What was it in our feedback condition that brought about the reduction in JDS and KSS while improving performance appraisals? The most optimistic explanation would suggest a genuine feedback effect: the driver’s awareness of their own functional state, enhanced by the OAMS, may have enabled them to apply their own means of adjusting their level of alertness. Alternative explanations include placebo effects and simple re-alerting effects of any display change that draws the driver’s attention in addition to their routine tasking. The Hawthorn effect can be discounted as our participants were always aware of the fact they were being monitored – both in feedback and no-feedback conditions.

The re-alerting mechanisms of such feedback have yet to be examined. Is it the continuous provision of feedback that makes the difference, or an invitation for an update action that the driver gets when the feedback screen goes blank after several minutes of low JDS readings? Our field study had no capacity to answer this question in full, as it afforded no control over the driver action in respect of the feedback display. For this question to be addressed, future studies will have to enforce greater control both over the mode of feedback provision and over the driver response. However, our analysis of the dynamics of JDS scores across individual driving periods offer a preliminary answer. We counted the number of JDS peaks in each driving period and measured each peak’s duration. If the observed difference in the average drowsiness levels (mean JDS scores over a driving period) was due to a genuine feedback effect then, according to biofeedback models (36) JDS peak duration should be shorter. Our findings confirm that expectation, suggesting that a genuine feedback mechanism was likely at play in reducing drowsiness under the feedback condition.

As is typical in field study contexts, the amount of control over participants action was limited. We had no influence on when they drove and for how long. As a result our data were distributed unevenly across the complete 24-hour cycle, with a relatively small number of observations in late evening and early morning hours and no data coverage between 1 and 4 am. This is representative of non-operational driving patterns but requires caution in generalising our findings to continuous operations, such as those in the mining industry. Variation in drive duration was also quite substantial - from a few minutes to nearly three hours. We accounted for this by including drive duration as a co-variate in all our analyses. However, both circadian phase and drive duration are likely to moderate the effect of OAMS feedback, and as such would benefit from a more targeted analysis with a more deliberate manipulation of drive duration and timing. This analysis would help answer important questions such as: how does the feedback effect change with increasing drive duration? At what time of day is it most/least pronounced?

Confirming the preliminary conclusion about the feedback mechanism responsible for our main finding would require ruling out a placebo effect (the effect of mere awareness of being monitored). This will require additional experimentation that might include presenting numbers,
unrelated to eye blinking, that drivers believe reflect their drowsiness. Such experimentation seems worth pursuing in future research.

The effect of OAMS feedback on both drowsiness and driving performance ratings was robust and statistically significant, which supports, in principle, the capacity of OAMS to generate improvement in both drowsiness and performance as a stand-alone system. However, the relatively small size of this effect leaves open the question of the system’s practical utility as a behaviour change agent. The practical benefits that OAMS feedback confers still need to be examined against alternative means of driver state monitoring.

Given that the purpose of OAMS is to provide alerts at higher levels of driver drowsiness, observation of a feedback effect on drowsiness and performance measures at low levels of OAMS-measured drowsiness is promising. Future research should replicate the study over night-time hours (e.g., between 10pm and 4am) when drivers are expected to be most drowsy, and incorporate long drives, where monotony of the task would be expected to induce drowsiness. These conditions would more closely replicate those found in military continuous operations and allow investigation of the utility of OAMS as a behaviour change agent in a more dangerous and realistic context.

6. Acknowledgements

The authors thank Warren Lewis, Dana Michalski, Gary Coombs and Philip Jacques for assistance with data collection; John Stewien and Elaine Lee for their help with data processing; Andy Tucker for advice on JDS score interpretation and Justin Fidock for commenting on earlier drafts of this paper.

7. References


