Accident Analysis and Prevention 50 (2013) 341–350

Contents lists available at SciVerse ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Efficient driver drowsiness detection at moderate levels of drowsiness

Pia M. Forsman[a], Bryan J. Vila[a,b], Robert A. Short[c], Christopher G. Mott[d], Hans P.A. Van Dongen[a,*]

[a] Sleep and Performance Research Center, Washington State University, Spokane, WA, USA
[b] Department of Criminal Justice and Criminology, Washington State University, Spokane, WA, USA
[c] College of Nursing, Washington State University, Spokane, WA, USA
[d] Pulsar Informatics Inc., Seattle, WA, USA

ARTICLE INFO

Article history:
Received 28 February 2012
Received in revised form 20 April 2012
Accepted 2 May 2012

Keywords:
Circadian misalignment
Driving performance
Drowsy driving detection
High-fidelity driving simulator
Lateral lane deviation
Steering metrics

ABSTRACT

Previous research on driver drowsiness detection has focused primarily on lane deviation metrics and high levels of fatigue. The present research sought to develop a method for detecting driver drowsiness at more moderate levels of fatigue, well before accident risk is imminent. Eighty-seven different driver drowsiness detection metrics proposed in the literature were evaluated in two simulated shift work studies with high-fidelity simulator driving in a controlled laboratory environment. Twenty-nine participants were subjected to a night shift condition, which resulted in moderate levels of fatigue; 12 participants were in a day shift condition, which served as control. Ten simulated work days in the study design each included four 30-min driving sessions, during which participants drove a standardized scenario of rural highways. Ten straight and uneventful road segments in each driving session were designated to extract the 87 different driving metrics being evaluated. The dimensionality of the overall data set across all participants, all driving sessions and all road segments was reduced with principal component analysis, which revealed that there were two dominant dimensions: measures of steering wheel variability and measures of lateral lane position variability. The latter correlated most with an independent measure of fatigue, namely performance on a psychomotor vigilance test administered prior to each drive. We replicated our findings across eight curved road segments used for validation in each driving session. Furthermore, we showed that lateral lane position variability could be derived from measured changes in steering wheel angle through a transfer function, reflecting how steering wheel movements change vehicle heading in accordance with the forces acting on the vehicle and the road. This is important given that traditional video-based lane tracking technology is prone to data loss when lane markers are missing, when weather conditions are bad, or in darkness. Our research findings indicated that steering wheel variability provides a basis for developing a cost-effective and easy-to-install alternative technology for in-vehicle driver drowsiness detection at moderate levels of fatigue.

1. Introduction

Drowsy driving (or driving while sleepy or fatigued) is a main contributor to road crashes (National Transportation Safety Board, 1999), as corroborated by various sources of data. In Europe, up to 20% of all traffic accidents are believed to be due to driver drowsiness (AWAKE, 2002). In the U.S., falling asleep while driving causes at least 100,000 crashes annually; 40,000 lead to nonfatal injuries, and over 1500 result in fatal injuries (Royal, 2002). As many as 28% of polled U.S. drivers admit to nodding off at the wheel at least once (National Sleep Foundation, 2009). In light of these disconcerting statistics, countermeasures against drowsy driving have received increased attention during the last couple of decades (Dinges et al., 1998).

To safeguard against drowsy driving, carmakers are developing technologies to monitor car-based metrics of driving performance and warn the driver of impending drowsiness. Such technologies typically rely on the detection of lane departures, large lateral deviations within the lane, and/or cessation of steering corrections (Galley et al., 2009; Seko et al., 1986; Victor, 2009). Whether these technologies truly serve a preventive purpose by detecting drowsiness sufficiently early on, without the help of physiological measures of sleepiness recorded from drivers themselves (see Vadeby et al., 2010), has not been convincingly demonstrated. In particular, there is an ongoing need to develop tools for reliably detecting driver drowsiness at relatively moderate levels of drowsiness, so that drowsy driver crashes can be anticipated and avoided well in advance.

In a literature search we found that researchers have tested at least 87 different metrics of driving performance for their potential

* Corresponding author at: Sleep and Performance Research Center, Washington State University Spokane, PO Box 1495, Spokane, WA 99210-1495, USA.
Tel.: +1 509 358 7755; fax: +1 509 358 7810.
E-mail address: hvd@wsu.edu (H.P.A. Van Dongen).

0001-4575/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.aap.2012.05.005

usefulness in detecting driver drowsiness. The metrics we considered are described in Appendix A and are discussed in Berglund (2007), Boyle et al. (2008), Fagerberg (2004), King et al. (1998), Kircher et al. (2002), Mattsson (2007), Otmani et al. (2005), Pizza et al. (2008), Tijerina et al. (1999), and Wierwille et al. (1994). We found no published large-scale comparisons between the available metrics, but there are a few comparisons within subsets of selected metrics. Depending on the study, different metrics stand out as those potentially most sensitive to driver drowsiness. For example, Friedrichs and Yang (2010) compared 31 metrics of driving performance and found that among these metrics, the average steering angular velocity was the most sensitive. Sandberg et al. (2011a) compared 18 metrics and reported that variability of lateral velocity was the most sensitive. In a joint effort, Berglund (2007) and Mattsson (2007) compared a set of 17 metrics and found that a linear combination of steering wheel direction reversals, vehicle path deviations, and standard deviation of lateral position was most sensitive to driver drowsiness.

These examples illustrate that no consensus exists regarding which metric or combination of metrics would be the most sensitive to driver drowsiness. Moreover, given the multitude of available metrics, some degree of collinearity among them is to be expected. Using high-fidelity driving simulator data from drowsy participants and alert controls studied in a laboratory setting, we set out to examine collinearity among the 87 driving metrics we found in the literature. From this work, we were able to develop a new approach to detecting driver drowsiness at moderate levels of drowsiness, which is presented here.

2. Methods

We used data from two laboratory-based, high-fidelity driving simulator studies, referred to here as Study A (Van Dongen and Belenky, 2010; Van Dongen et al., 2011a) and Study B (Van Dongen et al., 2010). The design of these studies was very similar; they are therefore described together here.

2.1. Participants

The total dataset included data from $N = 41$ participants aged 22–39. Study A contributed 25 participants (mean age ± SD: 27.3 ± 5.5; 13 men, 12 women). Study B contributed 16 participants (mean age ± SD: 27.5 ± 5.6; all men).

Participant inclusion criteria were: good health (by physical examination, blood chemistry and questionnaires) and not a current smoker, good sleep (by baseline polysomnography, at-home actigraphy, at-home sleep diary and questionnaires), no shift work or transmeridian travel within one month of entering the study, valid driver's license, and not susceptible to simulator adaptation sickness (by structured, supervised test driving of the simulator). Participants gave written informed consent, and were compensated for their time. Both studies were approved by the Institutional Review Board of Washington State University.

2.2. Protocol

Both studies were strictly controlled laboratory studies. Participants lived inside the laboratory continuously for 14 days (Study A) or 16 days (Study B). In Study A, participants were randomized to either a night shift condition ($n = 13$) or a day shift condition ($n = 12$). Fig. 1 shows the schedules of the two conditions in Study A. In Study B, all participants were assigned to a night shift condition essentially equivalent to that of Study A.

In Study A, participants came to the laboratory at 09:00. The night shift condition began with a baseline day, which included three sessions, at 12:00, 15:00 and 18:00, to practice the test procedures (described below), and which contained nighttime sleep (time in bed (TIB): 22:00–08:00). Day two in the laboratory involved a nap opportunity (TIB: 15:00–20:00) to help transition to the night shift schedule. Participants where then subjected to 5 days of night shift, during which they had daytime sleep (TIB: 10:00–20:00) and took the performance tests (described below) at 21:00, 00:00, 03:00 and 06:00. After the 5-day shift work period, participants were given a 34-h restart break in the laboratory, which involved a break from testing and contained a nap opportunity (TIB: 10:00–15:00) to transition back to a daytime schedule, nighttime sleep (TIB: 22:00–08:00), and another nap opportunity (TIB: 15:00–20:00) to transition back to the night shift schedule. After the restart break, participants were subjected to another 5 days of night shift, identical to the first 5 days. The study ended with a nap opportunity (TIB: 10:00–15:00) to transition back to a daytime schedule, and a nighttime recovery sleep opportunity (TIB: 22:00–08:00). Participants left the laboratory at 14:00 on day 14. See Fig. 1 (top panel).

The day shift condition also began with a baseline day that included three sessions, at 12:00, 15:00 and 18:00, to practice the test procedures, and contained nighttime sleep (TIB: 22:00–08:00). Day two in the laboratory involved daytime wakefulness (no performance testing) and nighttime sleep (TIB: 22:00–08:00). Participants were then subjected to 5 days of day shift, during which they had nighttime sleep (TIB: 22:00–08:00) and took the performance tests at 09:00, 12:00, 15:00 and 18:00. After the 5-day shift work period, participants were given a 34-h restart break in the laboratory, which involved a break from testing and contained two nighttime sleep opportunities (TIB: 22:00–08:00). After the restart break, participants were subjected to another 5 days of day shift, identical to the first 5 days. The study ended with another nighttime sleep opportunity (TIB: 22:00–08:00), and participants left the laboratory at 14:00 on day 14. See Fig. 1 (bottom panel). Total TIB and the total number of performance tests were identical for the night shift and day shift conditions.

In Study B, there was only a night shift condition, which was equivalent to that of Study A – except that the baseline and restart periods were each a day longer, both adding a nighttime sleep period (TIB: 22:00–08:00) and a daytime waking period without testing. The total number of performance tests was identical to that in Study A.

2.3. Measurements

During the two 5-day shift periods, sessions with performance testing were scheduled four times per day (time points 1–4) – see Fig. 1. Each session included a 10-min psychomotor vigilance test (PVT; Dinges and Powell, 1985); a 30-min driving session on a high-fidelity driving simulator; another 10-min PVT; and a brief (less than 15-min) neurobehavioral test battery, which included computerized versions of the Karolinska Sleepiness Scale (KSS; Åkerstedt and Gillberg, 1990), a visual analog scale of mood (Monk, 1989), the Positive and Negative Affect Schedule (Watson et al., 1988), a digit-symbol substitution task (Wechsler, 1981), performance and effort rating scales (Dinges et al., 1992), and a cardinal direction decision task (Gunzelmann et al., 2004). Thus, each driving session was paired with independent, established indices of fatigue (e.g., Belenky et al., 2003; Van Dongen et al., 2003, 2011a). The laboratory accommodated four participants at a time, and there were two driving simulators. Therefore, participants were randomly assigned to consistently either do the driving, preceded and followed by the PVT, first and the neurobehavioral testing second, or the other way around. Either way, each session had a 45-min break between PVT/driving/PVT and the neurobehavioral test battery.