Accident Analysis and Prevention 84 (2015) 134–143

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Drowsy drivers' under-performance in lateral control: How much is too much? Using an integrated measure of lateral control to quantify safe lateral driving

Roald J. van Loon [a,b,*], Rino F.T. Brouwer [a], Marieke H. Martens [a,b]

[a] The Netherlands Organisation for Applied Scientific Research TNO, Soesterberg, The Netherlands
[b] University of Twente, Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 11 May 2015
Received in revised form 27 July 2015
Accepted 13 August 2015
Available online 10 September 2015

Keywords:
Driver behaviour
Driver drowsiness
Fusing behavioural indicators
Safety thresholds
Real-time measurements

ABSTRACT

Internationally, drowsy driving is associated with around 20% of all crashes. Despite the development of different detection methods, driver drowsiness remains a disconcerting public health issue. Detection methods can estimate drowsiness by directly measuring the physiology of the driver, or they can measure the effect that drowsiness has on the state of the vehicle due to the behavioural changes that drowsiness elicits in the driver. The latter has the benefit that it could measure the net effect that drowsiness has on driving performance which links to the actual safety risk. Fusing multiple sources of driving performance indicators like lane position and steering wheel metrics in order to detect drowsiness has recently gained increased attention. However, not much research has been conducted with regard to using integrated measures to detect increased drowsiness within an individual driver. Different levels of drowsiness are also rarely classified in terms of safe or unsafe. In the present study, we attempt to slowly induce drowsiness using a monotonous driving task in a simulator, and fuse lane position and steering wheel angle data into a single measure for lateral control performance. We argue that this measure is applicable in real-time detection systems, and quantitatively link it to different levels of drowsiness by validating it to two established drowsiness metrics (KSS and PERCLOS). Using level of drowsiness as a surrogate for safety we are then able to set simple criteria for safe and unsafe lateral control performance, based on individual driving behaviour.

1. Introduction

Sleepiness at the wheel, commonly referred to as driver drowsiness or drowsy driving, is a disconcerting public health issue. Although different sources quote different statistics, it is safe to say that internationally around 20% of all crashes are somehow related to drowsiness (AWAKE, 2002; MacLean et al., 2003; Klauer et al., 2006). As such, there is an extensive amount of literature addressing the issue of drowsiness detection. Different methods to detect drowsiness-induced impaired driving have been established, and many of them attempt to assess the level of drowsiness by monitoring the physiology of the driver. Self-reporting methods, such as the Karolinska Sleeping Scale (KSS; Åkerstedt and Gillberg, 1990), are widely used to subjectively assess driver drowsiness and produce reasonable results in drowsiness-related studies (Kaida et al., 2006). Another prominent and less intrusive method was based on the finding that drowsiness goes along with slow eye movements (Erwin et al., 1973) including slow eye closures (Skipper and Wierwille, 1986). This resulted in the development of the PERCLOS (percentage of eye closure) P80 measure which uses a vision system to quantify drowsiness as the proportion of time in a given interval that the eyelids are at least 80% closed (Wierwille et al., 1994; Wierwille and Ellsworth, 1994; Dinges and Grace, 1998). Another method to estimate driver drowsiness is using the effect that drowsiness has on the state of the vehicle due to the behavioural changes it elicits. This has the advantage that it directly measures the net effect of drowsiness on driver performance. Deriving from the amount of literature on this topic, two aspects of vehicle state are key in detecting driver drowsiness: lane position related (e.g., Hanowski et al., 2008) and steering wheel related (e.g., McDonald et al., 2012). In an attempt to investigate the impact of roadside monotony on crash causation, Thiffault and Bergeron (2003) conducted a driving simulator study and concluded that drowsiness causes drivers to respond slower to lane

* Corresponding author at: The Netherlands Organisation for Applied Scientific Research TNO, Soesterberg, The Netherlands.
E-mail addresses: roaldvanloon@gmail.com, roald.vanloon@tno.nl (R.J. van Loon).

http://dx.doi.org/10.1016/j.aap.2015.08.012
0001-4575/

position deviations and with larger steering wheel movements. It so appears that drowsiness affects the driver's lateral control of the vehicle, and it has recently been argued that decreased lateral control performance can be used to not only detect imminent drowsiness, but also moderate levels of drowsiness (Forsman et al., 2013). This has the potential of using the level of drowsiness as a surrogate measure of safety to qualify lateral control performance as either safe or unsafe, because drowsiness measures like PERCLOS and KSS have known criteria (e.g., Hanowski et al., 2008; Sommer and Golz, 2010). However, as of yet no attempt has been made to do so on an individual level. In the present study, we use a high-fidelity driving simulator with a monotonous (night time) driving task in an attempt to fuse lane position and steering wheel measures into a single measure of individual lateral control performance, and set criteria for safe lateral control based on the PERCLOS and KSS drowsiness estimates. The result is a single vehicle-based real-time measure that quantitatively tells us if and how much lateral control performance degrades, independent of individual behavioural differences.

2. Methods

2.1. Participants

Seventeen participants (6 female and 11 male), age ranging between 28 and 56 years, participated in the experiment. Each participant was an experienced driver with at least eight years in possession of a driver's license and with a minimum mileage of 12,000 km per year. All participants were moderately to highly experienced with driving at night.

2.2. Driving simulator

The driving simulator used in this experiment was an modified vehicle mounted on a moving base, with a 180° view screen in front of the car, and displays placed behind the car in line with the rear-view mirrors (see Fig. 1). Many parameters including speed, pedal positions, lateral position, and steering wheel angle were recorded. The room was dark and the projection was adjusted to night-time driving with headlights visible (see Fig. 2). The road as well as the scenery were kept as monotonous as possible. Filters were added to the projectors in order to block blue light (470 nm), avoiding an excessive increased stimulation of the circadian mechanism (Brainard et al., 1988, 2001). A SmartEye eye-tracking system was installed in the simulator.

2.3. Roadway

A 9-km stretch of two-lane highway with various small radii curves as well as straight sections was repeated a number of times, such that the participant did not notice that he or she was driving the exact same road multiple times. The resulting road had no exits or on-ramps, no signs and there was no traffic present. A 2-km stretch of road was prepended for practice purposes.

Fig. 1. Driving simulator.

Fig. 2. Night time driving task, front view with visible headlights.

2.4. Procedure

Participants were asked to abstain from drinking alcohol and caffeine 24 h prior to the experiment. The experiment took place during daytime. Before being seated in the simulator, a questionnaire was completed containing various questions to check for physical fitness, including the onset and duration of sleep the night before as well as the onset and duration of sleep on average. Participants who consumed caffeine or alcohol, or participants who had less sleep the night before than on average were excluded from participation. This way all participants started the experiment in relatively the same physical fitness, i.e., not fatigued. Participants were also familiarized with the KSS scale. After completing the questionnaire, the participant was seated in the driving simulator for the experiment to begin. Upon completion, the participant was brought to a bright room and was offered some water, and stayed there at least 15 min until he or she was recovered enough to safely drive home.

2.5. Task

Participants were instructed to drive 100 km/h and to stay in the right lane. Every 5 min, the KSS was queried using a pre-recorded monotonous and low-pitched voice. The participant was continuously monitored by the experimenter from the control room, which had an one-way intercom system from the simulator room, and answers to the KSS were recorded by the experimenter. After 1 h, the experimenter stopped the experiment. In case the participant was obviously extremely tired or felt asleep, the experimenter stopped the experiment before the 1-h mark.

2.6. Data analysis

For each session signals from the driving simulator like lateral position and steering wheel movements were extracted. Additional parameters were calculated, most significantly the iteration of the repeated road segment and type of road segment (curve left, curve right, straight). Using the data from the SmartEye system the eyes open ratio was calculated and blinks were removed. All data (driving simulator, eyes open ratio, and KSS) were synchronized and the data from the stretch of highway used for practice purposes was removed.

Most drivers, at least in driving simulator studies, keep to a position off the lane centre (Pilutti and Ulsoy, 1999). Lateral performance can either be assessed as the driver's ability to maintain a specific path, e.g., lateral precision, or as the driver's ability to track the lane centre, e.g., lateral bias (Land and Horwood, 1995; Liu et al., 2009). The latter differs between individuals, but the former is a measure of lateral performance at the individual level. Because we want to measure at the individual level, we calculated lateral position using a 30-s rolling window of mean lateral position as the "zero point" lateral position. For the remainder of this paper, lateral position values will be the deviation from this zero point: negative lateral position values indicate driving to the left (and positive values to the right) of the driver's preferred lane position in a 30-s window. To make our results usable for real-time systems, we used a right-aligned rolling window (i.e., a rolling window over the preceding 30 s of data). PERCLOS was than calculated from the eyes open ratio using the same window size (30 s) and alignment (right-aligned).

We started by analysing whether the drowsiness manipulation using the monotonous night-time driving task had the predicted effect, that is whether participants truly became drowsy during the trip. First, we did this by exploring the PERCLOS and KSS measures over time. Next, we analysed the effects drowsiness had on lateral position and steering wheel angle. Due to vehicle momentum characteristics, it is possible that steering wheel behaviour and lateral position are moderated by changes in the speed of the vehicle (Wewerinke and Hogema, 2003). We therefore also made sure participants maintained the 100 km/h speed that was requested of them, and that any variations in speed were not dependent on level of drowsiness. Because steering wheel angle and lateral position are physically as well as behaviourally interdependent (behaviourally because steering wheel angle can be seen as a function of current lateral position or driver's estimated future lateral position), we are interested in the effect drowsiness has on the composite of these two. For now, we will focus on current lateral position. We initially differentiated between the lowest and highest levels of drowsiness based on thresholds from existing literature. For KSS, this was done by removing the median KSS score (KSS score of 5), and assign scores below the median as *low drowsiness* (which includes absence of drowsiness) and scores above the median as *high drowsiness* (Sommer and Golz, 2010). Similarly for PERCLOS, a value below 12.5% was considered to indicate *low drowsiness* and a value above 25% *high drowsiness* (Hanowski et al., 2008).

Using the results of drowsiness on lateral position and steering wheel angle, we moved on to develop an integrated measure of lateral control performance calibrated for each participant using a baseline of that same measure. Because the level of drowsiness is lowest at the start of the trip, we created this baseline using data from the start of the trip. Here, low drowsiness is when either PERCLOS or KSS is considered low, and high drowsiness is were either PERCLOS or KSS is considered high. Because steering behaviour and hence lateral control performance is dependent on road curvature, we also studied the effects with regard to curves. Participants drove the exact same stretch of road for several times, which gave us the possibility to compare subsequent repetitions. Finally, we validated our measure against known criteria for KSS and PERCLOS and set our own criteria accordingly. For this, we again differentiated between low and high levels of drowsiness, but also included moderate levels of drowsiness.

3. Results

In this section, we will first present exploratory results (Sections 3.1 and 3.2) and move on to develop the lateral performance measure based on lane position and steering wheel measures in Sections 3.3 and 3.4, which will then be validated against drowsiness measures in Section 3.5.

3.1. Drowsiness manipulation

To investigate whether the drowsiness manipulation was successful and had an effect within the 1 h trip, PERCLOS and KSS were averaged in 5-min bins and plotted in Fig. 3.

PERCLOS and KSS increased over time, indicating a successful drowsiness manipulation. PERCLOS measurements and KSS scores stabilized around the half hour, and PERCLOS was lower at the end. This is due to the fact that not all participants completed the full hour trip. One participant fell asleep just before the half hour mark, and an additional two fell asleep within 15 min after the half hour mark. The other participants completed the trip, but although every one of them indicated high drowsiness on the KSS scale, some of them never reached high PERCLOS measurements which causes the average PERCLOS to be lower and the interquartile ranges larger near the full hour mark compared to earlier in the trip. This shows us that time-in-trip across participants is not usable as a predictor for drowsiness because of individual differences, but we do have a wide range of PERCLOS and KSS scores for at least most of the participants, which opens the possibility to explore individual differences further.

3.2. Behavioural indicators

Before we are able to fuse lateral position and steering wheel measures into an indicator for safe lateral control performance using drowsiness as surrogate for safety, we need to know if and how drowsiness has an effect on lateral position and steering behaviour in the present study. As shown in Fig. 4, standard deviation in lateral position increased significantly at higher KSS scores across participants, $F(8,75) = 1.7$, $p < .05$, $\eta^2 = .15$, and also correlated significantly with PERCLOS, $r = .31$, $p < .01$.

As shown in Fig. 5, standard deviation of steering wheel angle also increased significantly at higher KSS scores across participants, $F(8,75) = 3.3$, $p < .01$, $\eta^2 = .25$. Also PERCLOS correlated significantly with steering wheel angle, $r = .22$, $p < .01$.

Overall average speed was 103.3 km/h (SD = 5.35), which is in line with what was asked of the participants (100 km/h). As shown in Fig. 6, speed did not differ between different levels of KSS, $F(8,75) = .40$, $p = \text{n.s.}$, and did not correlate with PERCLOS, $r = .02$, $p = \text{n.s.}$, indicating that drowsiness did not have an effect on speed.

3.3. Fusing lateral position and steering wheel angle

We found that drowsiness occurs within most trips (see Section 3.1) and that lateral position as well as steering wheel angle are affected by drowsiness (see Section 3.2). To fuse lateral position and steering wheel angle, we first analyse the effect of drowsiness on a composite of these two. Steering wheel angle is plotted as a function of lateral position for low and high drowsiness groups in Fig. 7 (based on KSS) and Fig. 8 (based on PERCLOS). This way, we can differentiate between corrective and non-corrective steering: that is, corrective steering is defined as steering to the right when driving to the left of the preferred lateral position, or steering to the left when driving to the right of the preferred lateral position. Consequently, the upper left and lower right quadrants in Figs. 7 and 8 indicate corrective as well as non-corrective steering, and the other two quadrants indicate solely non-corrective steering.

For low levels of drowsiness as well as high levels of drowsiness, 95% of the steering events are approximately evenly distributed across all quadrants. This makes distinguishing small corrective steering events from steering entropy difficult within the 95% confidence bounds. However, when looking at 99% of the steering events,

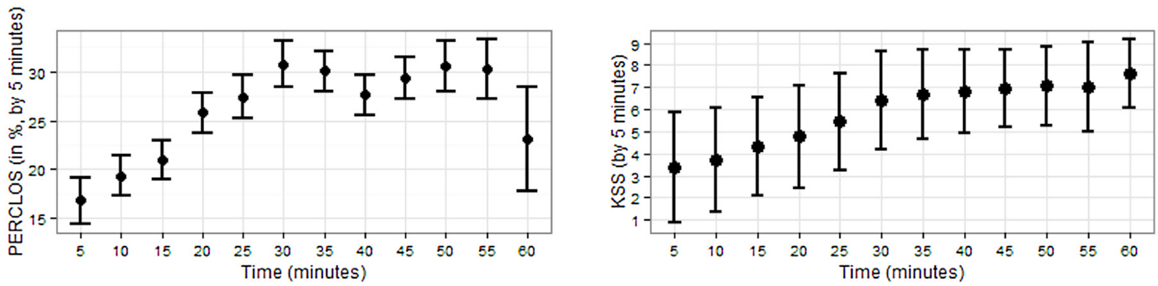

Fig. 3. Means and standard errors for PERCLOS (left) and KSS (right) over time for all trips.

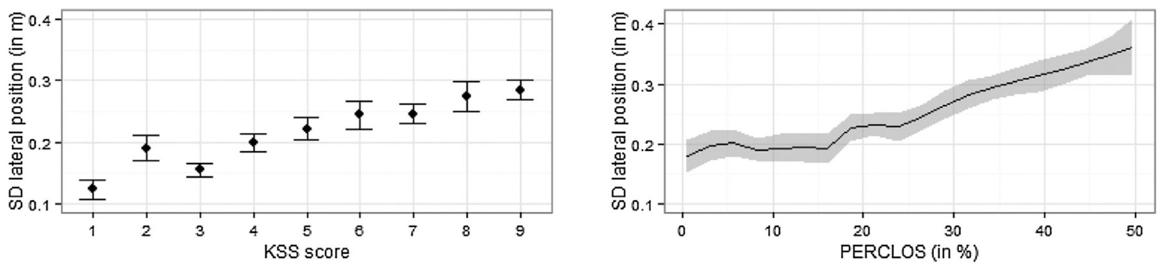

Fig. 4. Means and standard errors for standard deviation of lateral position, plotted for each level of KSS score (left) and by PERCLOS (right).

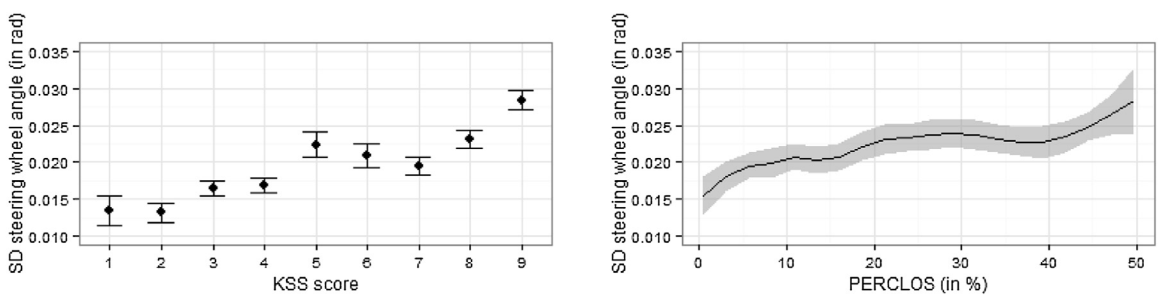

Fig. 5. Means and standard errors for standard deviation of steering wheel angle, plotted for each level of KSS score (left) and by PERCLOS (right).

the ratio of corrective steering events increases. We also see that for high levels of drowsiness this distinction becomes more apparent, indicating that drivers exhibit more severe steering wheel angles when lateral position deviates further from the driver's mean. Consequently, the proportion and size of corrective steering events appears to be an indicator of drowsy driving.

From these results we might be able to fuse lateral position and steering wheel angle into a measure that could detect drowsy driving based on corrective steering events, especially using the data outside the 95% confidence interval. However, because we do not always know the population (and hence confidence intervals) of steering events beforehand (for instance, in a real-time detection system), we need to add more weight to corrections that occur later and at lateral position further away from the driver's mean lateral position. This way the relative impact of steering events and lateral positions outside the 95% confidence interval increases. We propose a simple measure w as a function of the proportion of corrections r multiplied by a severity s over a given time interval S, stated as: $W = r \times S$. Here, the proportion of corrections r can be expressed as the number of corrective steering events divided by the number of all steering events in the same time interval. Severity s is the product of the maximum absolute lateral position p and the maximum absolute steering wheel angle a, stated as: $s = \max_{t \in S}\{|p_t|\} \times \max_{t \in S}\{|a_t|\}$.

The time interval S needs to encompass at minimum the time a driver needs to correct the steering wheel angle based on current lateral position (e.g., reaction time), both in a non-drowsy as well as a drowsy state. It also needs to be large enough to capture enough

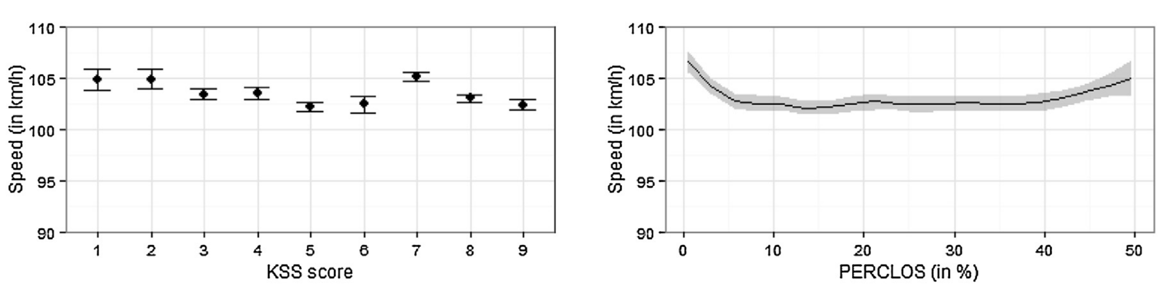

Fig. 6. Means and standard errors for speed, plotted for each level of KSS score (left) and by PERCLOS (right).

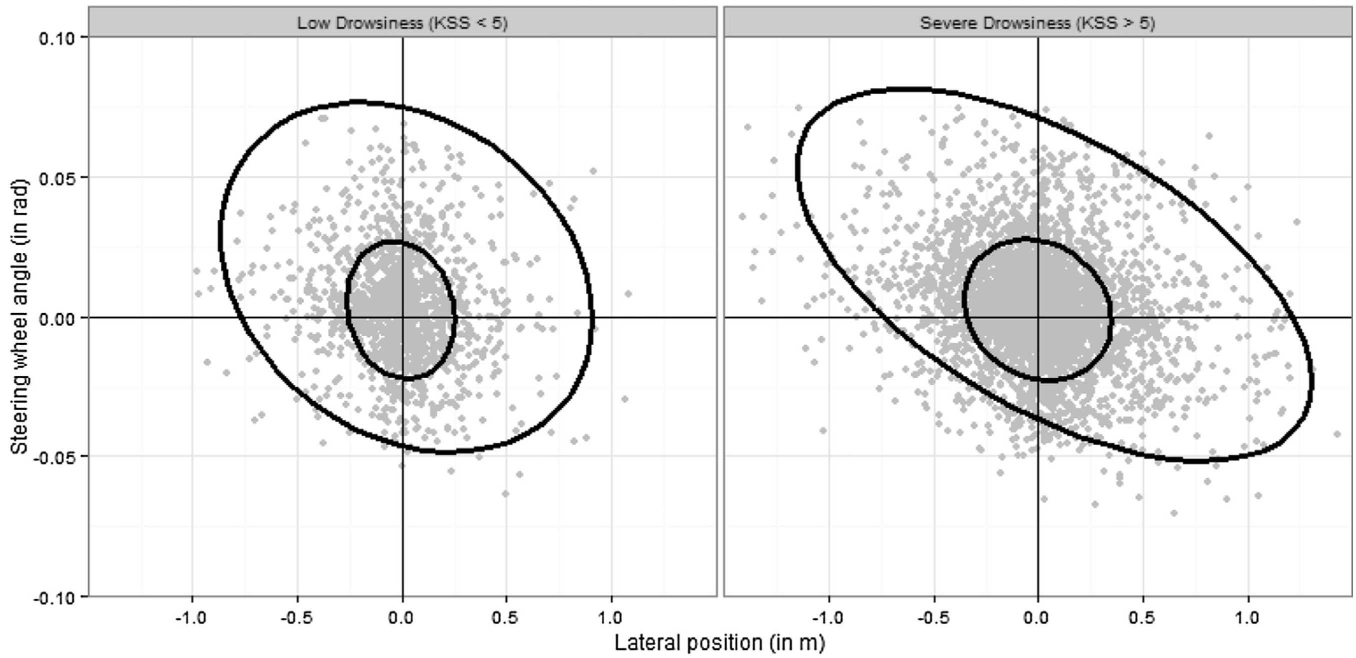

Fig. 7. Steering wheel angle (positive means steering to the right) as a function of lateral position, grouped by level of drowsiness (cf. KSS). Ellipses approximately contain 95% (inner ellipses) and 99% (outer ellipses) of data points.

steering events that are both corrective as well as non-corrective and to be less sensitive to sudden steering events. Finally, vehicle yaw rate is a function of time and hence changes in lateral position due to corrective steering are not instantaneous. For now, we pick an arbitrary time interval of 30 s, $|S| = 30$.

For the trip of a participant that became drowsy, this results in an increasing w as displayed in Fig. 9. We see that this participant has an increase in PERCLOS and crosses 12.5% (upper threshold for low level of drowsiness) around 1500 s into the trip. However w starts to increase earlier into the trip, at 600 s, which might indicate that w has a higher sensitivity towards the same construct that PERCLOS measures (i.e., drowsiness).

To be able to compare w between participants, a baseline w_b was calculated for each participant based on the first 2 min of data when driving a straight road section, and W was calculated as changes in w compared to this baseline, stated as: $W = w_b/w$. Changes in the resulting W, when considered a measure of lateral control performance, consequently indicate whether lateral control performance degrades (below 1.0) or improves (above 1.0) compared to the start of the measurement.

Within each trip, the same road was repeated approximately ten times, depending on full completion of the hour one trip. W is significantly different between low and high drowsiness across repetitions, $F(1,14) = 5.2$, $p < .05$, with W reaching lower values in later repetitions when drowsiness was high, $F(1,14) = 3.5$, $p < .05$, but not when drowsiness was low, $F(1,12) = .3$, p = n.s. (see Fig. 10). This indicates that W is sensitive to drowsiness while remaining insensitive to driving time when level of drowsiness remains low.

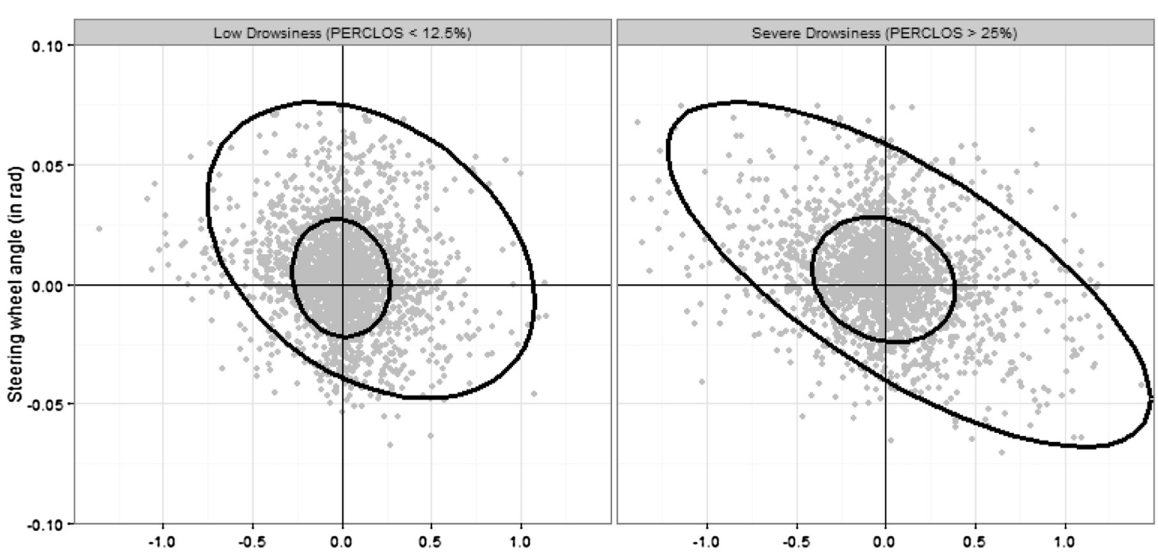

Fig. 8. Steering wheel angles (positive means steering to the right) as a function of lateral position, grouped by level of drowsiness (cf. PERCLOS). Ellipses approximately contain 95% (inner ellipses) and 99% (outer ellipses) of data points.

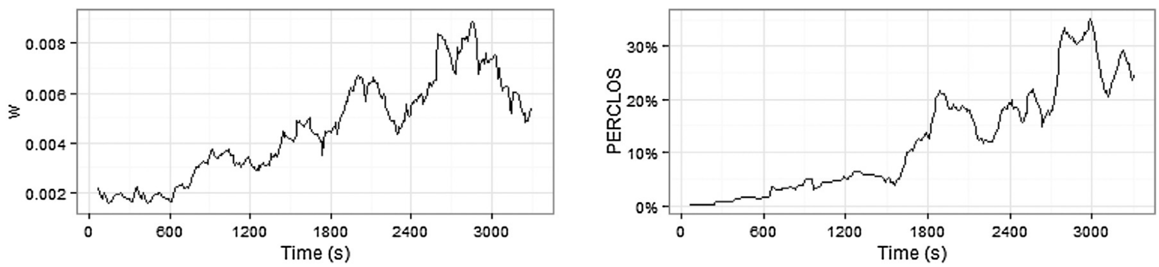

Fig. 9. Calculated w (left) and respective PERCLOS (right) for a single trip of one participant.

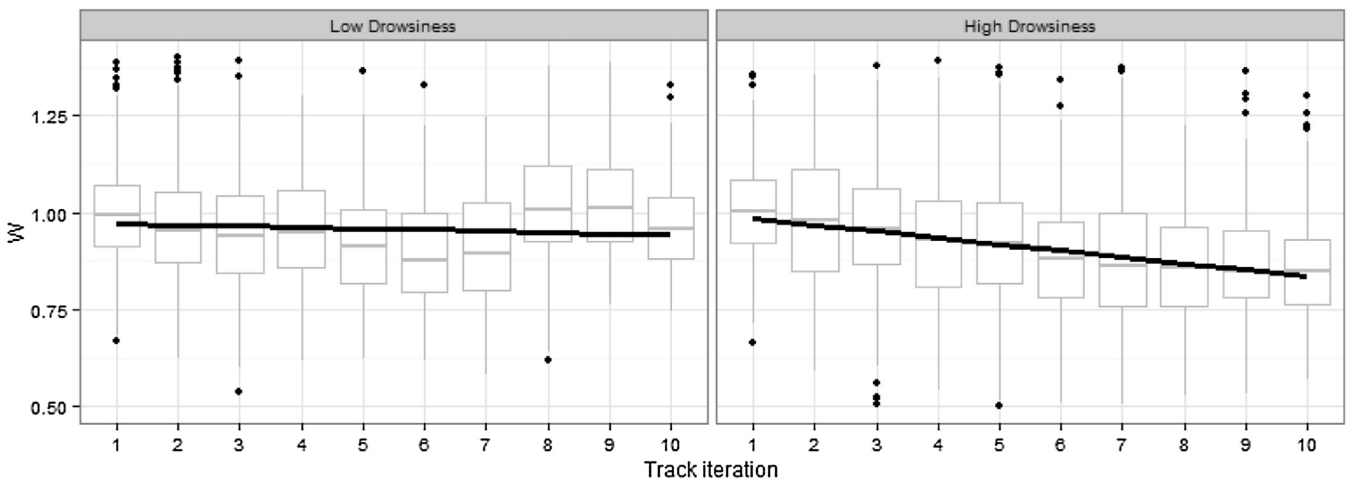

Fig. 10. Levels of W of all participants for low and high levels of drowsiness, grouped by iteration.

3.4. The effect of road curvature

Steering behaviour is inherently dependent on road curvature, and consequently our measure of lateral control performance W is significantly different between left curves, right curves and straight sections, $F(2,32) = 4.2$, $p < .05$, $\eta^2 = .16$ (sphericity was violated, Mauchly $\chi^2(2) = .48$, $p < .01$, so a Greenhouse–Geisser correction was applied, $\varepsilon = .66$). This was mainly because W did not change over time in left curves, $r = .01$, $p =$ n.s., but did decrease in right curves, $r = -.11$, $p < .001$ and even more so in straight sections, $r = -.21$, $p < .001$ (see Fig. 11).

W does not differ significantly between left curves, right curves and straight sections across repetitions when the driver show high levels of drowsiness, $F(2,11) = 1.4$, $p =$ n.s., but does for low levels of drowsiness at alpha 10%, $F(2,9) = 3.6$, $p < .10$. This indicates for high levels of drowsiness, W is not dependent on curvature. For low levels of drowsiness, it appears that the significant difference between left curves, right curves and straight sections at alpha 10% is caused by an increasing W when driving in left curves (see Fig. 12). Therefore, W can measure a decrease in lateral performance when drowsiness is high independently of curvature, but cannot measure an increase when drowsiness is low.

3.5. Validation and setting criteria

We found that a decrease in W occurs over time for high levels of drowsiness, but not for low levels of drowsiness. To validate whether W can be used as a measure of drowsiness, we need to make sure specific levels of W correlate with specific levels of drowsiness. Across participants, W is significantly dependent on KSS, $F(8,41) = 11.1$, $p < .001$, $\eta^2 = .69$ and on type of road section, $F(2,41) = 4.3$, $p < .05$ (although this effect is much smaller, $\eta^2 = .17$).

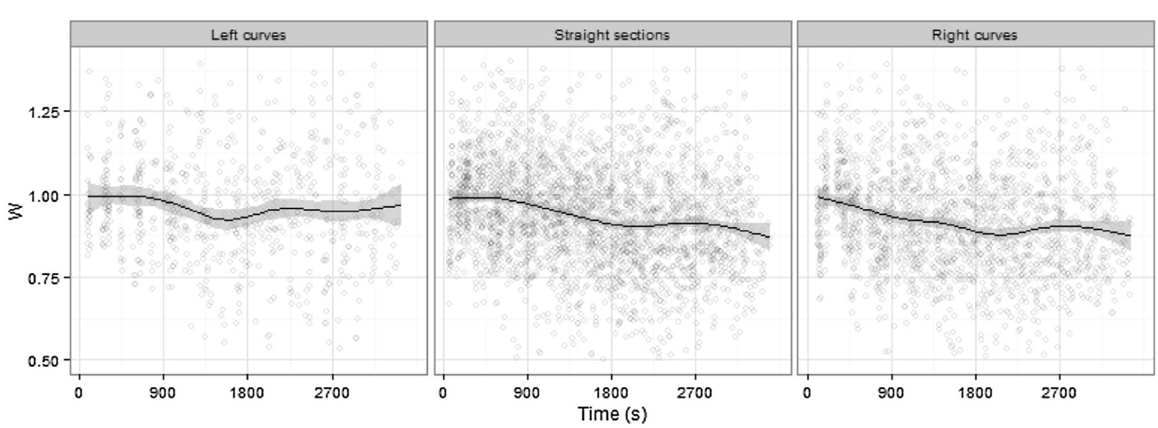

Fig. 11. Levels of W of all participants for all three types of road section over time. The line is a fitted spline of the mean with standard error.

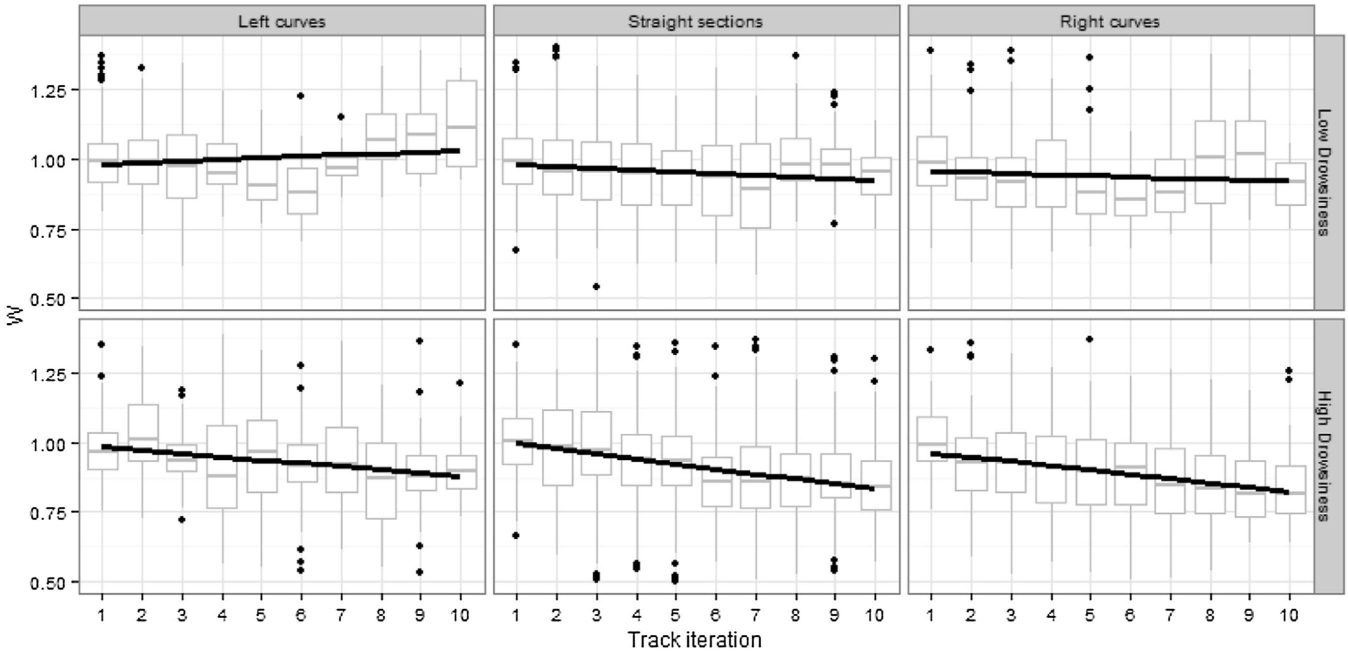

Fig. 12. Levels of *W* of all participants for all three types of road sections by iteration, by low level of drowsiness (top row) as well as high level of drowsiness (bottom row) and with a fitted regression line.

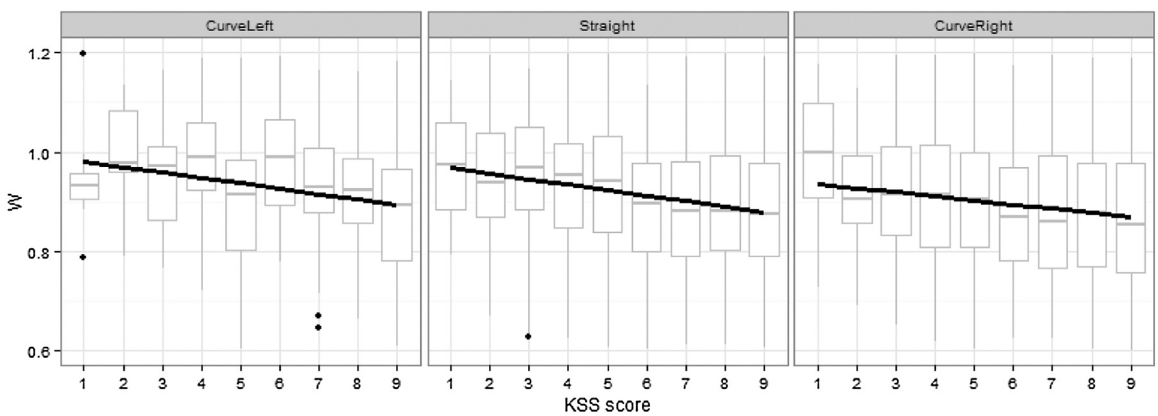

Fig. 13. Levels of *W* of all participants for all three types of road sections by KSS score, with a fitted regression line.

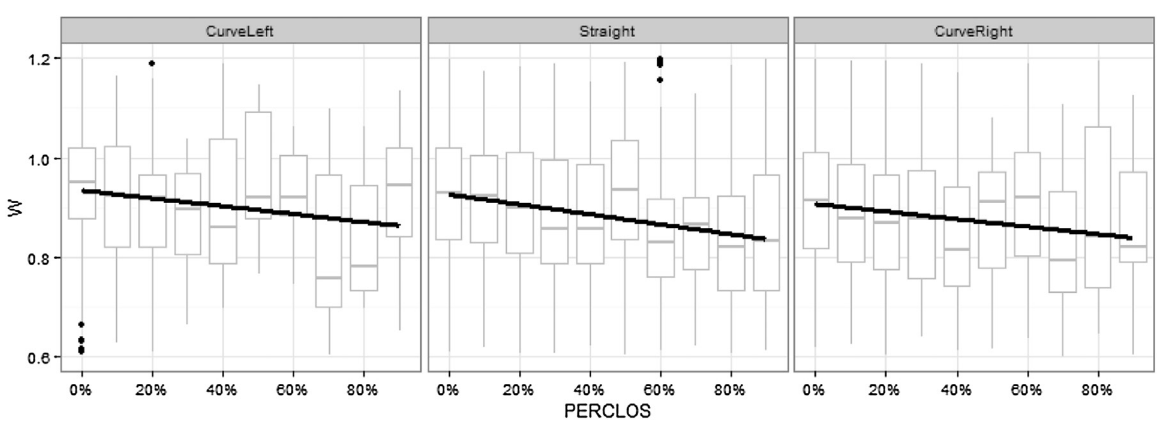

Fig. 14. Levels of *W* of all participants for all three types of road sections by PERCLOS, with a fitted regression line.

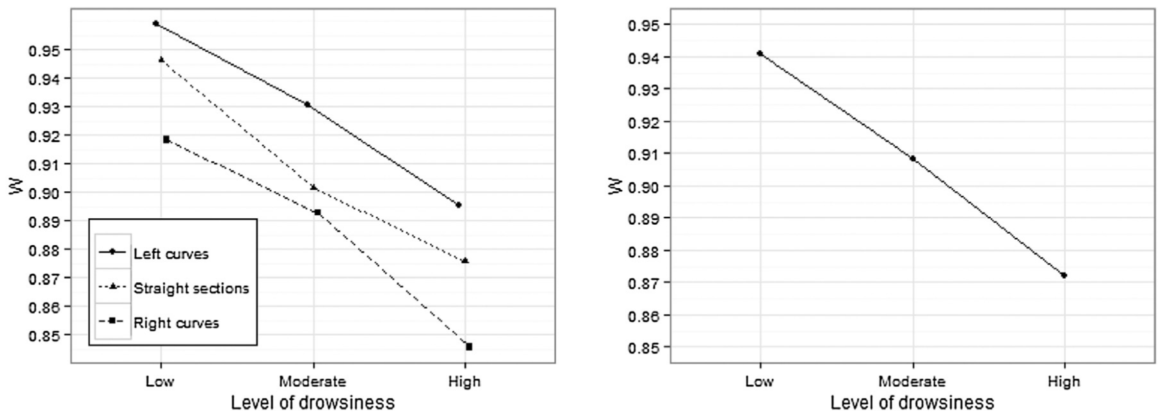

Fig. 15. Average W across all participants for low, moderate and high levels of drowsiness, by type of road sections (left) and averages across types (right).

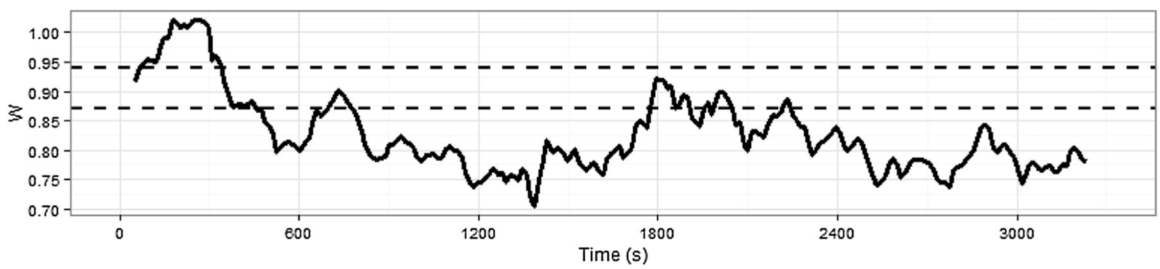

Fig. 16. Slightly smoothed W for a single participant over time (same participant as in Fig. 9), with the two criteria W = .94 (upper dashed line) and W = .87 (lower dashed line).

There is no interaction between these two, $F(16,41) = 1.8$, p = n.s. When KSS is considered nominally, it correlates significantly with W with low levels of W co-varying with higher KSS score, $r = -.20$, $p < .001$ (see Fig. 13).

W also significantly depends on PERCLOS across participants, $F(1,12) = 17.2$, $p < .01$, $\eta^2 = .59$, but not on type of road section, $F(2,12) = 1.0$, p = n.s. (see Fig. 14). Similarly to KSS, PERCLOS correlates negatively with W, $r = -.20$, $p < .001$.

When considering level of drowsiness as a surrogate measure of safe lateral control, we can set criteria for safe and unsafe lateral control using W. Average values of W where KSS and PERCLOS indicate low and high drowsiness were calculated, and values between low and high drowsiness were designated moderate drowsiness. Using these values, we can determine safe lateral control performance to be situations where W has not yet dropped below levels associated with low drowsiness (W = .94), and unsafe safe lateral control performance to be situations where W has dropped below levels associated with severe drowsiness (W = .87, see Fig. 15).

Plotting W for the same participant as in Fig. 9 (see Section 3.3) with these two criteria produces the graph in Fig. 16. We see that W for this participant decreases over time and crosses the W = .87 threshold around 600 s into the trip, indicating that this participant starts to exhibit too unsafe lateral driving. It also seems to recover slightly just after 600 s, but it never recovers to a safe level of lateral driving (W = .94).

4. Discussion

In this paper, we describe a new vehicle-based measure that is capable of real-time distinguishing between safe and unsafe lateral driving. As a surrogate for the level of safety of lateral driving, we used different validated measures for identifying levels of drowsiness from KSS self-reporting scores and PERCLOS measurements. Because we aimed for a vehicle-based measure capable of quantifying safe lateral driving on an individual level, it was important that we constructed this measure based on data that contained different levels of drowsiness within each participant, including baseline (non-drowsy) driving. The drowsiness manipulation (a monotonous highway driving task at night) that we applied turned out to be effective. Many drivers reached higher levels of drowsiness (as determined by higher KSS and higher PERCLOS scores) during the 1-h trip compared to when the trip started, and drowsiness manifested itself relatively early in the trip. This is in line with nearly all findings with regard to drowsiness when driving on monotonous roads and in monotonous environments (cf. Thiffault and Bergeron, 2003), and when driving in night time conditions on highways (Åkerstedt et al., 1994). For the development of the vehicle-based measure, we used lane position and steering wheel metrics. Steering wheel angle correlated significantly and positively with KSS scores and PERCLOS measurements, and is therefore subject to the effects of the drowsiness manipulation. A similar link between drowsiness and data from steering wheel angles has been found earlier in studies using KSS (Krajewski et al., 2009) and PERCLOS (McDonald et al., 2012) in an effort to develop drowsiness-predicting algorithms. Standard deviation from the driver's preferred position (that is, lateral position as a measure of lane keeping performance instead of lane position preference, cf. Liu et al., 2009) showed strong correlations with KSS as well as PERCLOS. Standard deviation of lateral position has been previously linked to both KSS (Ingre et al., 2006) and PERCLOS (Dingus et al., 1985), with similar results.

Drowsiness seems to have a particular effect on the severity of the steering wheel angles and the timing of steering events when considered a function of lateral position. For higher levels of drowsiness, drivers resort to more severe steering wheel angles when lateral position deviates further from the driver's mean. Corrective steering behaviour, that is, linking steering wheel angles to lateral position in real-time in an attempt to uncover the timing factors of the underlying behaviour, has not often been studied earlier. Notably, Thiffault and Bergeron (2003) found similar effects of drowsiness on steering wheel angles, although they did not study their data in real-time (but rather in 5 min blocks) and could not include measures related to lateral position in their analysis due to problems inherent with their experimental design. More specifically, no studies to our knowledge have attempted to include the effects of different levels of drowsiness on the changes in corrective steering behaviour. In the present study, we did not only find that steering wheel angles increase with drowsiness in general, but that this is mostly limited to those steering events that are corrective and occur further away from the driver's mean lateral position. This means that extreme values of steering wheel angle and lateral position are particularly indicative for driver drowsiness. The measure W developed in this study is based on this finding and has shown to be successful in differentiating between different levels of drowsiness within individuals, with W reaching lower values when drowsiness is more severe. This indicates that a decrease in W is quantitatively linked to decreased lateral performance.

Steering behaviour and lateral position are directly influenced by road curvature. Because W is a composite measure of these two, we investigated whether W could differentiate between different levels of drowsiness independently of curvature. For low levels of drowsiness, there was a small but marginally significant difference in W when comparing between left curves, right curves and straight sections. However for high levels of drowsiness, the decrease in W was independent of curvature. As a result, a decrease in W is more strongly associated with a higher level of drowsiness than an increase in W with lower levels of drowsiness. Apparently an increase in drowsiness seems to negate the influence of road curvature on W. This seems to be mostly limited to driving in left curves however. A possible explanation for this is the experimental setup: participants were asked to drive in the right lane in a two lane highway. Non-drowsy participants likely accept later and more severe steering wheel angles in the left direction when driving in left curves, because there is an emergency lane on the right hand side. Hypothetically, the rationale behind this acceptance is rather conscious (that is, consciously knowing it is safe to deviate to the right into the emergency lane) and might disappear when participants become more drowsy, resulting in similar steering behaviour for both left and right curves as well as on straight sections. In short, drowsy drivers might adjust their safety margins for lateral driving to a more generic model instead of curvature-specific. This effect has not been studied before and therefore lacks evidence, so more research is required to test this hypothesis. Nonetheless, more drowsy driving is always associated with a decrease in W. When presuming level of drowsiness to be an indicator for level of safety, it appears that the effect of road curvature on W disappears when lateral driving becomes more unsafe.

The data indicates that the W measure can be used to differentiate between low and high levels of drowsiness although it is merely constructed based on vehicle measures, and a validation was conducted to see if W correlates with drowsiness measures. We found that W correlates with KSS and also with PERCLOS. Low levels of W co-varies with high scores on KSS and high PERCLOS measurements, and this was independent of road curvature. This supports our earlier finding that a decrease in W is a valid method to estimate level of safety, as an increase in drowsiness can be considered unsafe. Additionally, we suggested specific levels of W as simple cut-off values where drowsiness was still low or absent ($W > .94$) and were drowsiness was severe ($W < .87$), which can be used as an indication of unsafe lateral control performance.

Although W is calibrated for an individual driver and requires limited calibration, using these criteria in support systems is not trivial due to several notable limitations. First and foremost, an accurate non-drowsy baseline measurement could be impeded when drivers start the trip when they are already drowsy. This does not affect the characteristics of W however, as a declining W over time still indicates unsafe lateral control performance. Secondly, speed remained stable in this experiment, and this is likely to be (at least partially) caused by the experimental setup in which we asked participants to drive a fixed speed (100 km/h). In reality however, speed might be more volatile as a result of driver's coping mechanisms with drowsiness, for instance by intentionally fluctuating speed by moving their foot on and off the accelerator pedal. The lack of traffic in our study could also negatively impact the external validity of our study, because traffic likely has an effect on speed, lateral position and also steering wheel angles. We do not know if and how these limitations affect W. Additionally, our results are valid for roads with high speeds like highways only (where drowsiness most likely occurs), but it is not known how W performs on urban roads. Future research could investigate these limitations more elaborately. Thirdly, a breach of the safety criteria could occur momentarily and does not necessarily mean the driver's lateral control performance is too unsafe to continue driving. Again, a more longitudinal measure, such as a declination in W or a time-below-threshold criterion, could be a suitable method to detect a decline in safety. Finally, although we know the characteristics of W in conjunction with slow impairment mechanisms like drowsiness, its applicability to detect unsafe driving due to short-term impairments (e.g., such as impairment due to temporary distraction), is not yet known. This would be a next step in order to develop W as an all-round measure able to detect long-term as well as sudden safety-critical changes in lateral control performance.

References

Åkerstedt, T., Czeisler, C.A., Dinges, D.F., Horne, J.A., 1994. Accidents and sleepiness: a consensus statement from the International Conference on Work Hours, Sleepiness and Accidents, Stockholm, 8–10 September 1994. J. Sleep Res. 3 (4), 195, http://dx.doi.org/10.1111/j.1365-2869.1994.tb00131.x.

Åkerstedt, T., Gillberg, M., 1990. Subjective and objective sleepiness in the active individual. Int. J. Neurosci. 52 (1–2), 29–37, http://dx.doi.org/10.3109/00207459008994241.

AWAKE, 2002. System for Effective Assessment of Driver Vigilance and Warning According to Traffic Risk Estimation. European Union Project IST-2000-28062, Retrieved from http://cordis.europa.eu/project/rcn/60961_en.html

Brainard, G.C., Hanifin, J.P., Greeson, J.M., Byrne, B., Glickman, G., Gerner, E., Rollag, M.D., 2001. Action spectrum for melatonin regulation in humans: evidence for a novel circadian photoreceptor. J. Neurosci. 21 (16), 6405–6412.

Brainard, G.C., Lewy, A.J., Menaker, M., Fredrickson, R.H., Miller, L.S., Weleber, R.G., . . ., Hudson, D., 1988. Dose–response relationship between light irradiance and the suppression of plasma melatonin in human volunteers. Brain Res. 454 (1–2), 212–218, http://dx.doi.org/10.1016/0006-8993(88)90820-7.

Dinges, D., Grace, R., 1998. PERCLOS: A Valid Psychophysiological Measure of Alertness as Assessed by Psychomotor Vigilance (TechBrief NHTSA No. FHWA-MCRT-98-006).

Dingus, T.A., Hardee, L.H., Wierwille, W.W., 1985. Detection of Drowsy and Intoxicated Drivers Based on Highway Driving Performance Measures (IEOR Department Report No. 8492). Virginia Tech, Blacksburg, VA.

Erwin, C.W., Volow, M.R., Gray, B., 1973. Psychophysiologic Indices of Drowsiness. Society of Automotive Engineers, New York.

Forsman, P.M., Vila, B.J., Short, R.A., Mott, C.G., Van Dongen, H.P.A., 2013. Efficient driver drowsiness detection at moderate levels of drowsiness. Accid. Anal. Prev. 50, 341–350, http://dx.doi.org/10.1016/j.aap.2012.05.005.

Hanowski, R.J., Bowman, D., Alden, A., Wierwille, W.W., Carroll, R., 2008. PERCLOS+: Moving Beyond Single-Metric Drowsiness Monitors (No. 2008-01-2692). SAE International, Warrendale, PA, Retrieved from http://www.sae.org/technical/papers/2008-01-2692

Ingre, M., Åkerstedt, T., Peters, B., Anund, A., Kecklund, G., Pickles, A., 2006. Subjective sleepiness and accident risk avoiding the ecological fallacy. J. Sleep Res. 15 (2), 142–148, http://dx.doi.org/10.1111/j.1365-2869.2006.00517.x.

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., Fukasawa, K., 2006. Validation of the Karolinska sleepiness scale against performance and EEG variables. Clin. Neurophysiol. 117 (7), 1574–1581, http://dx.doi.org/10.1016/j.clinph.2006.03.011.

Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J., 2006. The Impact of Driver Inattention on Near Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data (No. DOT HS 810 594). National Highway Traffic Safety, Washington, DC.

Krajewski, J., Golz, M., Sommer, D., 2009. Detecting sleepy drivers by pattern recognition based analysis of steering wheel behaviour. In: Presented at the Der Mensch im Mittelpunkt technischer Systeme, Dusseldorf.

Land, M., Horwood, J., 1995. Which parts of the road guide steering? Nature 377 (6547), 339–340, http://dx.doi.org/10.1038/377339a0.
Liu, C.C., Hosking, S.G., Lenné, M.G., 2009. Predicting driver drowsiness using vehicle measures: recent insights and future challenges. J. Saf. Res. 40 (4), 239–245, http://dx.doi.org/10.1016/j.jsr.2009.04.005.
MacLean, A.W., Davies, D.R., Thiele, K., 2003. The hazards and prevention of driving while sleepy. Sleep Med. Rev. 7 (6), 507–521, http://dx.doi.org/10.1016/S1087-0792(03)90004-9.
McDonald, A.D., Schwarz, C., Lee, J.D., Brown, T.L., 2012. Real-time detection of drowsiness related lane departures using steering wheel angle. Proc. Human Factors Ergonom. Soc. Annu. Meet. 56 (1), 2201–2205, http://dx.doi.org/10.1177/1071181312561464.
Pilutti, T., Ulsoy, A.G., 1999. Identification of driver state for lane-keeping tasks. IEEE Trans. Syst. Man Cybernet. – Part A: Syst. Hum. 29 (5), 486–502, http://dx.doi.org/10.1109/3468.784175.
Skipper, J.H., Wierwille, W.W., 1986. Drowsy driver detection using discriminant analysis. Hum. Fact. 28 (5), 527–540, http://dx.doi.org/10.1177/001872088602800503.
Sommer, D., Golz, M., 2010. Evaluation of PERCLOS based current fatigue monitoring technologies. IEEE, 4456–4459, http://dx.doi.org/10.1109/IEMBS.2010.5625960.
Thiffault, P., Bergeron, J., 2003. Monotony of road environment and driver fatigue: a simulator study. Accid. Anal. Prev. 35 (3), 381–391, http://dx.doi.org/10.1016/S0001-4575(02)00014-3.
Wewerinke, P.H., Hogema, J.H., 2003. Modelling Driver Behaviour in Combined Longitudinal and Lateral Control (TNO-Report No. TM-03-D001). TNO Human Factors, Soesterberg, The Netherlands.
Wierwille, W.W., Ellsworth, L.A., 1994. Evaluation of driver drowsiness by trained raters. Accid. Anal. Prev. 26 (5), 571–581, http://dx.doi.org/10.1016/0001-4575(94)90019-1.
Wierwille, W.W., Ellsworth, L.A., Wreggit, S.S., Fairbanks, R.J., Kirn, C.L., 1994. Research on Vehicle-Based Driver Status/Performance Monitoring: Development, Validation, and Refinement of Algorithms for Detection of Driver Drowsiness (NHTSA Report No. DOT HS 808 247).