

THE 8TH INTERNATIONAL CONFERENCE ON
DRIVER DISTRACTION AND INATTENTION



DDI 2022 GOTHENBURG

Lindholmen Conference Centre & online
October 19–20, 2022

ABSTRACT BOOK

INDEX

1.1 Driver Monitoring I

- Katja Kircher4
Driver attention monitoring and visual sampling from relevant and irrelevant targets
- Alwyn Hoffman.....8
Automated real-time detection of truck driver non-compliance
- Dr. Nadja Schömig 11
Investigating different driver-in-the-loop strategies on driver's eye glances and intervention behavior in partial automated driving
- Adriano Palao..... 14
Euro NCAP presents a 2-stage method to assess Driver State Monitoring systems as part of its 5-star consumer test programme

1.2 The Disconnected Driver

- Thomas Streubel 17
How does trust in vehicle automation effect the hand-over process? A Wizard-of-Oz study on public roads
- Dan Boguslavsky21
Effect of engagement with a Trivia game on driver's sleepiness and behavioural adaptation in a partially automated vehicle
- Natasha Merat25
Do drivers engage in NDRTs when travelling in an AV in real-world motorway environments?
- Derek Albert.....28
Brief Online Mindfulness Training in Young Drivers: A Pilot Randomised Controlled Trial

2.1 Driver Monitoring II

- Gaëtan Perrotte34
Body posture and physiological indicators for drowsiness detection in a partial automated driving.
- Che-tsung Lin.....38
Data Augmentation via Neural-Style-Transfer for Driver Distraction Recognition
- Mingming Zhao.....42
AutoConduct: a novel dataset for in-vehicle driver posture monitoring

2.2 Understanding and Assessing Visual Attention

- Vasiliki Kondyli.....46
Analysing Driver (In)Attentiveness: Towards a Cognitive Complexity Model Combining Visuospatial and Interactional Parameters
- Katja Kircher50
Train driver attention is influenced by the type of railway signalling system
- Emma Tivesten.....53
The influence of alcohol and automation on drivers' visual behavior during test track driving
- Ibrahim Öztürk.....57
How does night-time driving and engagement in a cognitive distraction task affect detection of peripheral targets?

3.1 Safety and performance assessment

- Tuomo Kujala60
Key requirements and a method for measuring in-vehicle user interfaces' distraction potential
- Pierluigi Olleja64
Quantification of driver's side-glance frequency and duration in straight highway driving
- Jonas Bärghman67
Validation of an eyes-off-road crash causation model for virtual safety assessment
- Christer Ahlström71
Strategical use of peripheral vision in driving
- Daniel Trommler.....74
Evidence accumulation modeling for the Detection Response Task when combined with the Box Task

3.2 Drowsiness and Fatigue

- Assaf Botzer78
Estimating Driver Fatigue Based on Heart Activity and Respiration Rate
- Johanna Wörle.....85
What leads drivers to illicitly nap during conditionally automated driving?

Stephanie Dabic Degrand 89

Assessment of the effectiveness of several countermeasures in reducing driver fatigue and associated risks for safety during autonomous driving.

Markus Tomzig 94

Who benefits from napping in automated driving? – Effects of chronotype on subjective sleep inertia

Ke Lu 98

Heart rate variability as an indicator for driver fatigue, different effects of time of day and time-on-task

4.1 HMI Design and Countermeasures

Dr. Nadja Schömig 101

“I’m Sad When You’re Distracted” – Effectiveness and User Experience of an Innovative Driver Monitoring System for Partially Automated Driving

Annika Larsson 104

To have and to hold during L2: Hands on wheel keeps drivers in lane while mind on road speeds response

Lara Désiré 109

Attentional demands of using an application for real-time traffic information feedback in road operators’ vehicles

Arun Muthumani 113

Human-machine interface designs assisting drivers of automated vehicles during transitions: evaluation from an end-user perspective

4.2 Distraction from smart phones and other technology

Yi-ching Lee 118

Psychosocial Needs and Factors Contribute to Problematic Phone Use while Driving

Burcu Arslan 122

TXT N DRV: A Systematic Review of the Effect of Texting while Driving on Driver Performance

Vanessa Beanland 125

Distracted, unfocused and risky driving in adolescents: Associations with sleep issues, inattention and sluggish cognitive tempo symptoms

Jeremy Wolter 130

Managing the driver distraction risk posed by technology: Changes to Australian road laws aimed at driver behaviour

Poster pitch session

Claus Marberger 139

Non-driving related task engagement in highly automated vehicles: How to mitigate emerging motion sickness?

David Large 143

Deriving Extended Keystroke Level Model Resumability Operators: An Occlusion Study

Paul Hemeren 146

Biomotion Triggering Driver Attention for Cyclists at Night

Eléonore Henry 149

Effect of agency on the occurrence of car sickness: incidence for autonomous driving

Arthur Portron 153

Ocular activity reflecting mental models’ dynamics is influenced by autonomous driving duration

Tina Morgenstern 157

Age effects on the Box Task combined with a Detection Response Task

Madelaine Coffey 160

“Who’s Got the Remote Control?” Understanding Driver Distraction and Inattention in the Context of Teleoperation and the Passenger Experience

Katerina Bucsuházy 163

Different types of human failure causing crashes

Driver attention monitoring and visual sampling from relevant and irrelevant targets

Katja Kircher^{1*}, Christer Ahlström^{1,2}

¹ VTI, Linköping, Sweden

² Department of Biomedical Engineering, Linköping University, Linköping, Sweden

*katja.kircher@vti.se

Abstract: Driver attention is often assessed via glance behaviour, typically by measuring glances away from the forward roadway or by directly measuring glances to non-driving related targets. This approach can be used to detect distracting events, but it does not check whether all situationally relevant targets are sampled. Here, we evaluate the usefulness of the MiRA-theory as basis for attention assessment. A field study was conducted with 23 participants driving an instrumented vehicle on an urban route. The participants wore a head-mounted eye tracker. Data reduction included the identification of target areas that needed to be sampled, whether they were sampled or not, and whether relevant or irrelevant other traffic was present. Additionally, a gaze-by-gaze analysis identified gaze direction, purpose, and target. As predicted, drivers sampled all required target areas that necessitated a glance away from forward. Target areas roughly in the forward direction, like zebra crossings, were probably sampled with peripheral vision, but this could not be reliably confirmed with the equipment used. The glance direction distribution was found to correspond well to the a-priori-defined requirements. A higher number of parallel requirements induced a larger share of glances with the purpose to check for traffic. Relevant traffic was monitored more than irrelevant traffic. A higher number of parallel requirements was associated with reduced spare visual capacity. Nominal glance target identification was less linked to the requirements. We therefore recommend that “traditional” glance-based attention assessment should be complemented with a purpose-based glance assessment protocol coupled with situation dependent pre-defined requirements.

1. Introduction

The assessment of driver attention or distraction is often done by investigating visual sampling strategies. The two most used methods are eye tracking and visual occlusion, sometimes in combination (Kujala, Kircher, & Ahlström, 2021). Visual occlusion can inform us about the attentional demand and predictability of a given situation, and in combination with eye tracking it allows an assessment of which glances are necessary and which are “redundant” (Kircher, Kujala, & Ahlström, 2020). When determining distraction, a common procedure is to use eye tracking to measure what a driver looks at with foveal vision and then to classify those glances as “relevant for driving” or “not relevant for driving”. Distraction is then assumed to occur if a driver is found to foveate targets that are “not relevant for driving”, which is in line with several of the often-cited definitions of driver distraction (Regan, Hallett, & Gordon, 2011).

However, it is typically not considered whether drivers *miss* relevant information, except for the research body centring around hazard perception. Here, a common approach is to investigate if and when drivers detect specific pre-defined hazards (Samuel & Fisher, 2015). In one of the rare field studies, Kaya, Girgis, Hansma, and Donmez (2021) found that drivers frequently miss over-the-shoulder checks for cyclists when turning, especially when turning right. Otherwise, and especially in real world studies, it is unusual to investigate whether *all* relevant information in a given situation is sampled (Ahlström, Kircher, Nyström, & Wolfe, 2021). Apart from the focus on “wrong targets”, this can also

be due to the challenge in identifying all relevant targets a-priori.

The theory of Minimum Required Attention (MiRA) provides a framework to identify relevant target areas which actual glance behaviour can be compared to (Kircher & Ahlstrom, 2017). Within this framework, a driver is attentive if and only if all relevant information was sampled in a timely manner. Based on traffic regulations, road layout and intended manoeuvres, the relevant target areas and the so-called MiRA-zone, within which the sampling has to take place, can be identified.

Here we explore whether the MiRA theory can be applied to empirical data from a real-world setting. We combine the notion of being attentive (as in having sampled all relevant information) with the approach of classifying glances in a purpose-related fashion instead of by the physical glance targets per se (Ahlström et al., 2021). Based on the above, in combination with the fact that most journeys are collision-free, we assume that

1. most relevant information is usually sampled,
2. sampling patterns are situation dependent, and
3. situations with more attentional requirements leave less spare capacity.

2. Method

Twenty-three participants drove a 12 km long route within the town of Linköping, Sweden. They were equipped with a head-mounted eye tracker (SMI glasses 2.0, SensoMotoric Instruments, Teltow, Germany) and drove an instrumented Volvo V60. The instruction was to drive as one normally would. Traffic density varied with time of day.

2.1 Analysis

Four intersections with different priority regulations (priority, yield, stop, traffic light) were selected for analysis. The participants went straight ahead in all four intersections. For each such event it was determined which areas had to be sampled from which “zone” along the route. For example, in the stop-controlled intersection, it was necessary to check for traffic from the left and right in a specific zone, which started where the crossing road was no longer obscured by houses, and which ended before the intersection was entered. Also, two zebra crossings had to be sampled before they were crossed.

For each such zone the required glance direction was determined (left, right, forward, or behind). It was also noted whether several requirements had to be fulfilled at the same time, that is, whether several requirements were “active” in parallel. Such combinations could be equal across intersections – for example checking a zebra crossing in combination with the forward roadway occurred on several occasions. These similar occurrences were analysed together.

For each such requirement combination the presence of road users in the required areas was coded regardless of whether they were glanced at or not. “Traffic encounter types” were identified as collision course, potential collision course, no collision course, or no road user present. Traffic that was present but not in a required area was classified as “irrelevant”.

Each participant’s eye tracking data were assessed with respect to whether the required areas were sampled while in the corresponding zones. The glance purpose was estimated as well, assuming that a first glance towards a required area was to check for traffic from that direction, and that repeated glances to another road user in the area were to keep track of its trajectory. Glances in the forward direction to no particular target were categorised as “default” and glances towards traffic lights or similar were coded as “information collection”. Glances were also encoded based on the type of gaze target as in “traditional” glance analysis and with respect to their direction in relation to the direction of travel.

3. Results

All MiRA-requirements that necessitated a gaze away from the forward direction were fulfilled by glancing at least once in the required direction within the zone. Whether all zebra crossings were sampled is more difficult to tell based on foveal eye tracking data, with 43 confirmed and 72 unclear cases. No incidents where pedestrians were not given priority occurred, though.

The sampling direction distribution in the different requirement combinations clearly reflected the direction of the required areas. If required areas were present in several directions at the same time, the glance direction was shared across areas. The more required areas there were within a single zone, the lower the share of “default” glances, indicating that demands were higher in these situations. This is also reflected by the higher share of glances with the purpose to check for traffic (Figure 1).

All relevant traffic, that is, all road users present in required areas, were sampled at least once. In contrast, only about half of the irrelevant traffic (road users outside of required areas) were glanced at. If relevant traffic was present at the same time, only around 25 % of the irrelevant road users were glanced at.

The classification into target types showed that motorised traffic was the most frequent glance target for a combination of the required areas forward and to the right, whereas cyclists and other targets were glanced at frequently when required areas existed in the forward, left and right direction.

Grouping by interaction type instead, it turned out that pedestrians and cyclists were glanced at more frequently when traffic on collision course was present than when on potential or no collision course (Figure 2). Default glances are less frequent with “higher” levels of interaction.

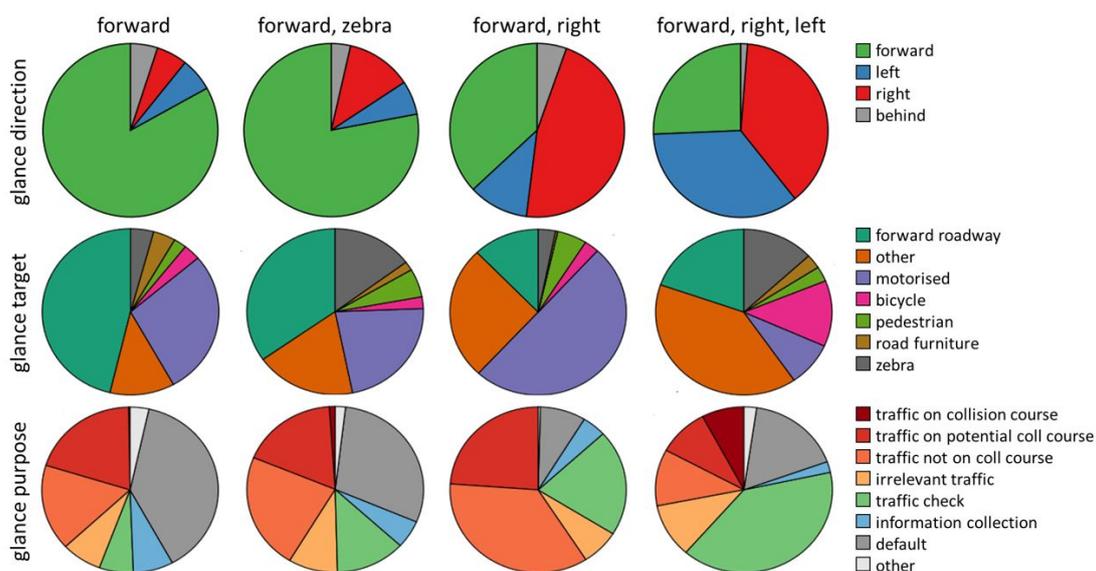


Figure 1. Glance direction, target and purpose, depending on different coinciding MiRA-requirements.

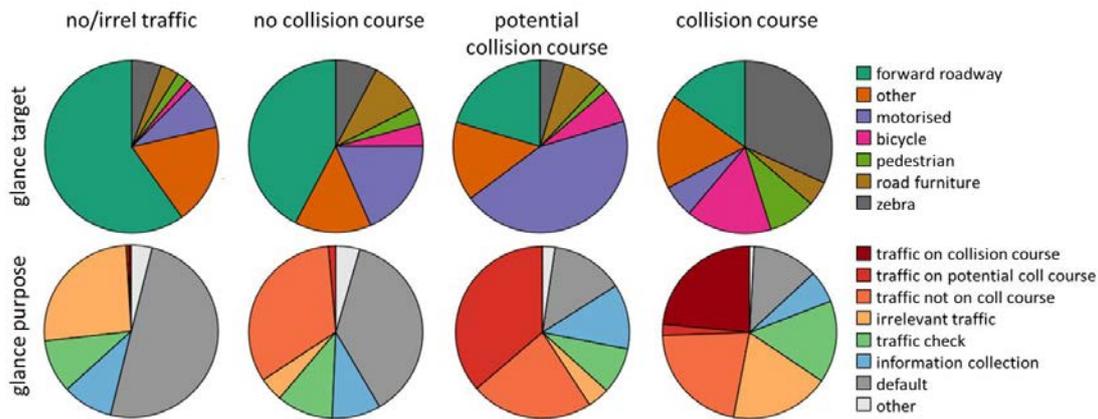


Figure 2. Glance target and purpose depending on traffic encounter type.

4. Discussion

A glance classification based on physical targets is likely to miss important aspects of driver attention for several reasons. Firstly, it is dependent on what happens to be there – in the current dataset cyclists and pedestrians were the road user groups that were most frequently on collision course with the participants, because their traffic light was green at the same time as the participant was turning right, across their (in this case rather busy) path. Also, the high frequency of “other” targets in the “forward, left and right” combination can be attributed to the fact that when checking for potential cross traffic the road was empty most of the time, which led to the foveated glance target essentially being “an empty road”, categorised as “other”.

If one considers glance purpose instead, the classification appears more logical. If several areas require attention simultaneously, the share of glances devoted to checking for traffic increases. Whether or not any traffic is present will then affect the following action. While this type of categorisation may appear less objective, it has been argued that this is not necessarily the case (Ahlström et al., 2021).

As hypothesised, the participants fulfilled all non-forward MiRA-requirements, with some methodological uncertainty for requirements in forward direction. Still, there were numerous glances to “other” targets. Due to the controlled nature of the study, these targets were not completely unrelated to driving. The finding still serves as an indication that a driver who glances at “other things” can be fully attentive – it may even be required to glance towards an area away from forward to be classified attentive, which occasionally result in glances to “other” targets.

When using the directional and purpose-based classification of glances, the behaviour is predictable and linked to requirements. Drivers meet the attentional requirements, and when they have capacity left, they spend it on “default” glances, taking in redundant information. Also, they invariably monitor relevant traffic, and if they have capacity left, they also look at irrelevant traffic.

Eye tracking gives access to foveal but not peripheral vision, which is a shortcoming that became especially apparent here when assessing sampling of zebra crossings.

For future research on driver attention, we suggest investigating further which glance behaviour can be said to fulfil an attentional criterion, instead of assuming that one glance in the corresponding direction is enough. We also recommend looking into the role of peripheral vision.

5. Conclusions

Empirical evidence from real-world driving shows that drivers’ glance behaviour follows the predictions made by the theory of Minimum Required Attention. Drivers adapt their gaze pattern to the situational requirements, dealing with relevant information first and foremost. If spare capacity is available, drivers can also sample redundant or irrelevant information, while still being attentive according to the MiRA theory.

6. Acknowledgments

This work was supported by the Swedish Strategic Vehicle Research and Innovation Programme (FFI; grant number 2017-05526).

References

- Ahlström, C., Kircher, K., Nyström, M., & Wolfe, B. (2021). Eye Tracking in Driver Attention Research—How Gaze Data Interpretations Influence What We Learn. *Frontiers in Neuroergonomics*, 2(34). doi:10.3389/fnrgo.2021.778043
- Kaya, N., Girgis, J., Hansma, B., & Donmez, B. (2021). Hey, watch where you’re going! An on-road study of driver scanning failures towards pedestrians and cyclists. *Accident Analysis & Prevention*, 162, 106380. doi:<https://doi.org/10.1016/j.aap.2021.106380>
- Kircher, K., & Ahlstrom, C. (2017). Minimum Required Attention: A human-centered approach to driver inattention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 59(3), 471-484. doi:10.1177/0018720816672756
- Kircher, K., Kujala, T., & Ahlström, C. (2020). On the difference between necessary and unnecessary glances away from the forward roadway: An occlusion study on the motorway. *Human Factors*, 62(7), 1117-1131. doi:10.1177/0018720819866946

- Kujala, T., Kircher, K., & Ahlström, C. (2021). A Review of Occlusion as a Tool to Assess Attentional Demand in Driving. *Human Factors*, 0(0), 00187208211010953.
doi:10.1177/00187208211010953
- Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis and Prevention*, 43(5), 1771-1781.
doi:10.1016/j.aap.2011.04.008
- Samuel, S., & Fisher, D. L. (2015). *Evaluation of the minimum forward roadway glance duration critical to latent hazard detection*. Paper presented at the Transportation Research Board 94th Annual Meeting, Washington DC, USA.
<http://trid.trb.org/view.aspx?id=1337244>

Automated real-time detection of truck driver non-compliance

Alwyn J. Hoffman^{1*}, Pierre Schutte¹, Schalk J. Rabè¹

¹ School of Electrical, Electronic and Computer Engineering, North-West University, South Africa

* alwyn.hoffman@nwu.ac.za

Abstract: Trucks contribute disproportionately to fatal traffic accidents. Overloading, speeding, and fatigue are primary causes of truck accidents. Speeding is an example of driver inattention while fatigue causes driver distraction due to sleepiness. Overloading is controlled by weigh-in-motion scales, but driver inattention can result in the evasion of such monitoring systems. Non-intrusive detection of these behavioural deviations can enable effective action against non-compliant drivers without causing disruption to compliant drivers. This paper describes a novel system that combines GPS tracking data with road monitoring data to automatically detect non-compliant behaviour in real time. Driver behaviour is compared against population behaviour based on incident statistics, thus separating compliant from non-compliant drivers. Results made available to roads authorities and fleet owners enable action against non-compliant drivers but with no impact on compliant drivers, thus minimising disruptions to economic activities.

1. Introduction

The importance of truck safety has been widely reported (Zhang, Yau, & Zhang, 2014) (Douglas, 2009) (Kemp, Kopp, & Kemp, 2013). In 1997, 98 percent of the fatalities from crashes between a truck and a passenger vehicle were occupants of the passenger vehicle (G.A.O., 1999). According to (Chen, Sieber, & Lincoln, 2015) commercial trucks were involved in 3,464 fatal, 73,000 injury-causing and 241,000 property-damaging crashes in 2012 in the US alone. They found that it would be beneficial to road safety if high-risk carriers receive reviews of driver compliance with safety regulations.

Mahaboon (Mahaboon, 2014) found that drivers reported high non-conformance with fatigue regulations, and that speeding violations was a strong predictor of crash involvement. Batool et al (Batool, Hussain, Kanwal, & Abid, 2018) found that long trucks are involved in fatal accidents because of drivers' risky driving behaviours, speeding and overloading. Another finding was that increased monitoring to enforce compliance with work hour rules resulted in increased unsafe driving practices due to speeding (Scott, Balthrop, & Miller, 2020).

The use of weigh-in-motion technology for overload control was proposed by Jacob et al (Jacob & Feypell-de La Beaumelle, 2010) to allow trucks to be weighed in traffic flow, without any disruptions to operations. In South Africa overloading is controlled through a combination of static and weigh-in-motion scales operated by the South African National Roads Agency (SANRAL) (SANRAL, 2017).

An IoT system for enhancing road safety proposed by Jabbar et al (Jabbar, Shinoy, & Kharbech, 2019) collects trip data, GPS coordinates, average and maximum speed and driving behavior for drivers' risk assessment and to detect extreme road user behavior.

This paper proposes a novel IoT solution that combines data from both road transport operators and the roads authority to detect non-compliance in real time, enabling effective action against offenders without negative implications for non-offenders.

2. Method

A pilot system was implemented to monitor compliance behaviour of freight trucks on the highway between Durban and Johannesburg. Data was collected from existing ANPR cameras and weigh-in-motion scales along the route, and from GPS tracking systems of 48 vehicles, over a period of 12 months. As the same vehicle is always driven by the same driver, incidents identified for a particular vehicle represent the behaviour of the corresponding driver. More than 5,000 trips were completed during this period, with each truck completing between 5 and 230 trips.

A list of incident types was defined to characterize unsafe road behaviour. To limit deployment cost incidents were restricted to those that could be detected using existing data. Driver inattention was represented by vehicles passing a WIM scale without using the WIM lane and speeding. Distraction incidents were based on non-compliance with fatigue regulations, resulting in drowsiness. The presence of each incident type could be identified using the above data sets.

ANOVA was applied to verify if truck driver identity was significantly related to incident rates. From the incident statistics per vehicle and for the population we calculated driver t-statistics per incident type to identify drivers displaying extreme behaviour.

3. Results

Table 1 below displays statistics for available observations and extracted incidents, including the total number and average number per vehicle. The results of the ANOVA analysis in Table 2 below confirms that vehicle ID is a significant indicator of incident prevalence, given that all F-statistics are much larger than 1 while all p-values are very small.

The histograms for t-statistics associated with each incident type are displayed in Figure 1. For WIM-scale incidents most drivers display a low incidence of misconduct, while there are several outliers with a much high incidence of misconduct compared to the population average. A significant fraction of drivers is involved in speeding, while a

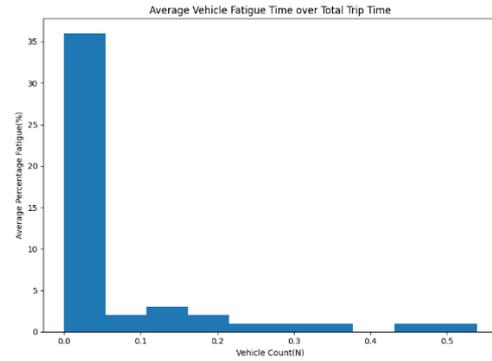
small percentage of drivers infringe significantly on fatigue regulations. The results obtained from the ANOVA analysis are thus confirmed by the t-statistics.

Table 1 Statistics on incidents recorded

| Statistic | WIM non-compliance | Speeding | Fatigue non-compliance |
|--|--------------------|----------|------------------------|
| Total number of observations | 10,718 | 8,047 | 8,047 |
| Average observations per vehicle | 282 | 168 | 168 |
| Total number of incidents | 1,552 | 7,770 | 298 |
| Average incidents per vehicle | 40.8 | 162 | 6.2 |
| Incidents as percentage of observation | 14.5% | 96.6% | 3.7% |

Table 2 ANOVA results assessing relationship between vehicle ID and incident rate

| Incident Type | F-statistic | p-value |
|------------------------|-------------|----------|
| WIM non-compliance | 8.13 | 2.99E-54 |
| Speeding | 60.5 | 5.7E-37 |
| Fatigue non-compliance | 4.56 | 7.55E-23 |

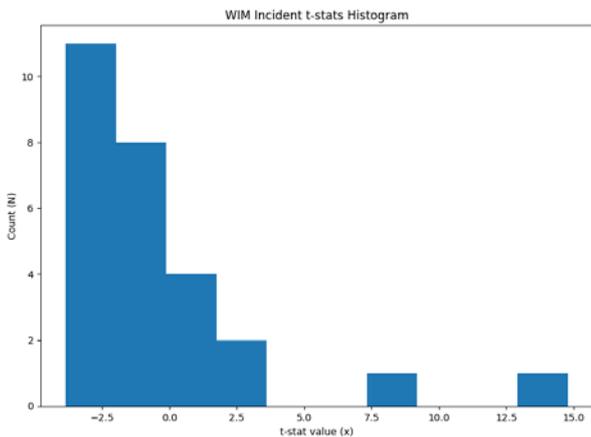


(c) Fatigue incidents

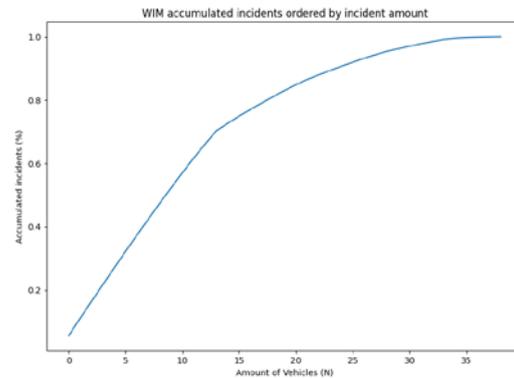
Figure 1 Histograms of driver incident rate t-statistics

4. Discussion

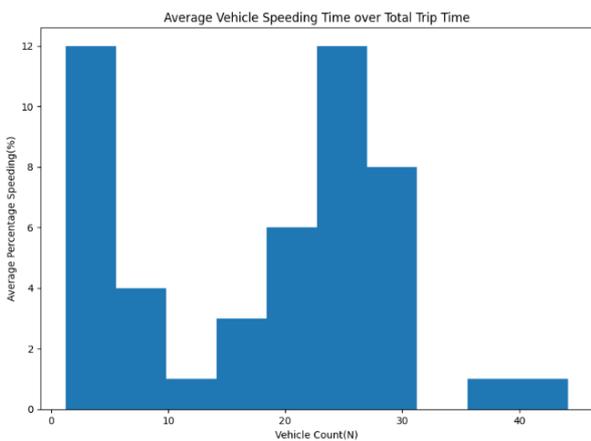
To practically reduce the prevalence of non-compliant behaviour it is necessary to act against perpetrators. This will be much less disruptive to cargo flows if action taken against a minority of truck drivers will eliminate most non-compliant behaviour. To investigate if a minority of truck drivers cause most non-compliant behaviour, we ranked drivers based on number of incidents observed and calculated the fraction of incidents represented by an increasing fraction of drivers, starting with those drivers that committed the most offenses. Figure 2 shows that by acting against 33% of drivers it is possible to eliminate 66% of WIM scale, 84% of speeding and 92% of fatigue offenses.



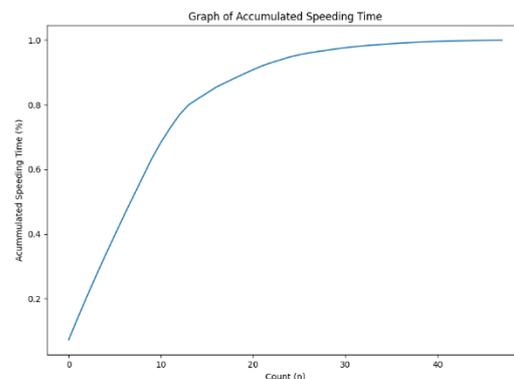
(a) WIM scale incidents



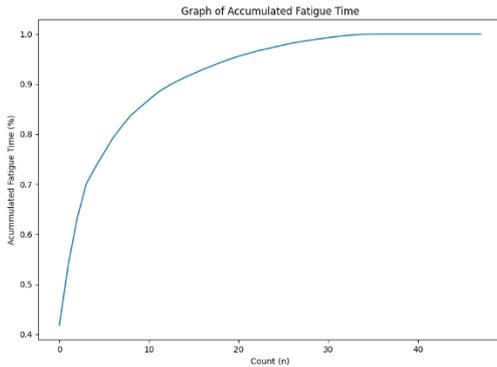
a) WIM scale incidents



(b) Speeding incidents



b) Speeding incidents



c) Fatigue incidents

Figure 2 Accumulative fraction of incidents caused by accumulative number of ranked drivers

5. Conclusions

We demonstrated that incidents reflecting truck driver inattention and distraction can be extracted by combining traffic data from different sources in real time. As the proposed concept does not require new hardware infrastructure it can be deployed at low cost. ANOVA provided evidence of a strong relationship between driver identity and non-compliance behaviour. t-statistics indicated that some drivers display extreme non-compliant behaviour. A minority of drivers was found to cause most of the non-compliance incidents. As the incident data is collected non-intrusively it is possible to apply enforcement against offenders without causing disruptions to compliant drivers.

Recommend future work should combine incident data with crash and insurance claims statistics to predict fatalities and insurance losses from observed non-compliant behaviour.

6. Acknowledgments

This work was funded by merSETA and SANRAL.

References

- Anonymous. (2021, July 28). *62% of South African road accidents involve trucks*. Retrieved from South Africa's Freight News: <https://www.freightnews.co.za/article/62-south-african-road-accidents-involve-trucks-0>
- Batool, I., Hussain, G., Kanwal, N., & Abid, M. (2018). Identifying the factors behind fatal and non-fatal road crashes: a case study of Lahore, Pakistan. *International Journal of Injury Control and Safety Promotion*, Vol. 25, No. 4, 401–407.
- Chen, G., Sieber, W., & Lincoln, J. (2015). NIOSH National Survey of Long-Haul Truck Drivers: Injury and Safety. *Accident Analysis & Prevention*, 85(1), 66–72.
- Douglas, M. (2009). *Commercial motor vehicle driver safety: an application of ethics theory*. Texas: University of North Texas.
- G.A.O. (1999). *TRUCK SAFETY: Motor Carriers Office Hampered by Limited Information on Causes of Crashes and Other Data Problems*. Washington, D.C.: United States General Accounting Office.

- Hijacking of motor vehicles*. (n.d.). (Statistics South Africa) Retrieved February 26, 2022, from <http://www.statssa.gov.za>
- IET Digital Library. (n.d.). *Author Guide - IET Research Journals*. Retrieved November 27, 2014, from IET Digital Library: <http://digital-library.theiet.org/journals/author-guide>
- Jabbar, R., Shinoy, M., & Kharbech, M. e. (2019). Urban Traffic Monitoring and Modeling System: An IoT Solution for Enhancing Road Safety. *International Conference on Internet of Things, Embedded Systems and Communications* (pp. 13-18). IEEE.
- Jacob, B., & Feypell-de La Beaumelle, V. (2010). Improving truck safety: Potential of weigh-in-motion technology. *IATSS Research* 34, 9-15.
- Kemp, E., Kopp, S., & Kemp, E. (2013). Six days on the road - Will I make it home safely tonight? Examining attitudes toward commercial transportation regulation and safety. *The International Journal of Logistics Management*, 24(2), 210-229.
- Mahaboon, J. (2014). *An Investigation into Factors Influencing Hazardous Materials Truck Crashes in Thailand*. Ph.D. Thesis. New South Wales, Australia: The University of New South Wales.
- SANRAL. (2017). *SANRAL Annual report 2016*.
- Scott, A., Balthrop, A., & Miller, J. (2020). Unintended responses to IT-enabled monitoring: The case of the electronic logging device mandate. *Journal of Operations Management*, 152-181.
- Zhang, G., Yau, K., & Zhang, X. (2014). Analyzing fault and severity in pedestrian–motor vehicle accidents in China. *Accident Analysis and Prevention*, 73, 141–150.

Investigating different driver-in-the-loop strategies on driver's eye glances and intervention behavior in partial automated driving

Nadja Schömig*, Katharina Wiedemann, Tristan Wehner, Alexandra Neukum

Würzburg Institute for Traffic Sciences (WIVW GmbH), Robert-Bosch-Str.4, 97209 Veitshöchheim, Germany

*schoemig@wivw.de

Abstract: The presented simulator study compared two different driver-in-the-loop strategies on driver's eye glances and intervention behavior at system limits in partial automated driving with a control condition without any strategy: A state-dependent strategy achieved by a driver monitoring system and a situation-dependent strategy by using a monitoring request. The results showed visible differences in gaze reaction times and intervention times. However, the effects were overlaid by strong individual differences in monitoring behavior during partial automated driving, so that none of these differences reached statistical significance. The qualitative analysis of single critical events indicated that those could be successfully avoided by the strategies. The subjective evaluations revealed that a monitoring request was perceived as more helpful in preparing for a system limit.

1. Introduction

When driving with a partial automated system the driver is still requested to monitor the driving environment and to be ready to react any time to a system limit or error. Two different strategies can support the driver in these tasks: A driver monitoring system (DMS) aims to keep the driver permanently *in* the loop (Merat et al., 2018). It observes driver's visual attention and triggers a warning if eyes are taken off from the driving task for too long. Such attention warnings proved to be successful in getting participants to monitor the road (see Victor et al., 2018; Blanco et al., 2015; Schömig & Kaussner, 2014). Another strategy is to bring the driver *back into* the loop only when necessary. A so-called monitoring request (MR) could ask the driver in uncertain situational circumstances to increase effort in monitoring to be better prepared in case the situation requests a driver intervention. The effectiveness of MR was previously investigated by Gold et al. (2013), Lu et al. (2019) and Louw et al. (2017a, b).

The presented study investigated which driver-in-the-loop (DIL) strategy leads to better intervention behavior at a system limit and compared it to a control condition without any strategy: A state-dependent strategy achieved by a DMS or a situation-dependent strategy via an MR.

2. Method

2.1 Test environment

The test was conducted in the WIVW driving simulator with motion system (see www.wivw.de). The simulator is equipped with an eye tracking system by SmartEye®.

2.2 Test scenarios and test course

The test course consists of a three-lane highway including 12 test scenarios (15 minutes duration). Each scenario first contains a phase where the driver drives with the partial automated system at a set speed of 100 km/h. In 8 of the 12 scenarios, after 45 sec driving an obstacle suddenly appears on the test vehicle's lane in 10 s distance

which is not detected by the system and therefore requires a driver intervention. The obstacle consists of safety beacons positioned across the complete width of the lane so that a complete lane change is necessary in order to avoid a collision. The scenarios differ in the direction of the lane change (depending on the initial lane when entering the scenario: either to the left or to the right), whether one or two lanes have to be changed (depending on the number of lanes blocked), and whether an additional braking maneuver is necessary because traffic is approaching from behind or not. The obstacle and traffic do not appear in the scene until 10 seconds before reaching the system limit. In the other 4 scenarios (so called distractor scenarios) the obstacle is either on a different lane or there is no obstacle at all so that no driver intervention is necessary. After having passed the obstacle the driver should stay on the target lane and drive manually until the end of the section. The 12 scenarios are arranged into a continuous driving course which can be driven through.

2.3 Automated system and HMI

A prototypical partial automated system requiring the driver to keep hands on the steering wheel was implemented in the driving simulator. It did not respond to the obstacle, but remained active until the driver deactivated the system by pressing a button at the steering wheel, braking or steering. A simple prototypical display in the instrument cluster was used to indicate the system status. In the active state, a green circle with the text "system active" was displayed, no status display was shown when the system was inactive.

2.4 Non-driving-related task (NDRT)

During the partial automated drive, drivers watched a video without dialogues, but with an acoustic background. It was presented on a tablet mounted at glove compartment level. Both, gaze and head direction towards the display were used as indicators for distraction. To encourage subjects to focus their attention on the video, even though this is not allowed in partial automated driving, they were explicitly instructed to do this for scientific purposes.

2.5 Driver-in-the-loop strategies

The study examined two different driver-in-the-loop strategies (DIL strategies) and compared them to a condition without any strategy. This factor was implemented as a between factor.

- State-dependent strategy (DMS condition): The drivers' gaze behavior was monitored via DMS. If they looked away from the road for too long (> 4 s), they received an Eyes-Off warning. This message was displayed directly on the video screen as long as the driver looked to it. It consisted of a text box with the information "Please monitor the traffic situation".

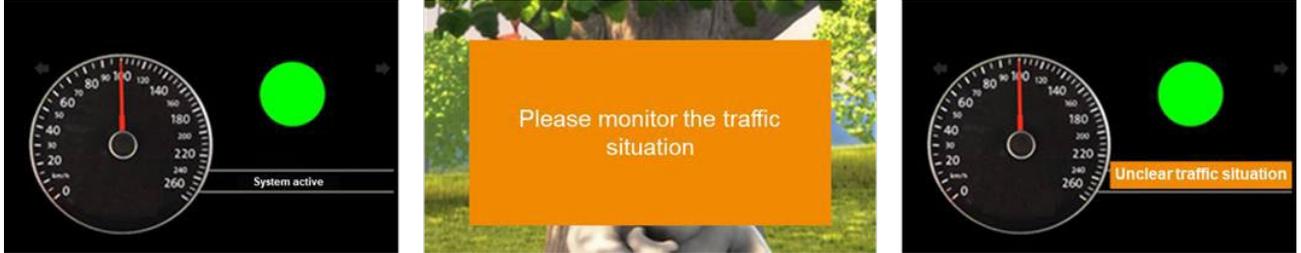


Fig. 1. HMI-Feedback for the different DIL strategy conditions. Left: Baseline condition Middle: DMS condition; Right: MR condition.

2.6 Test sample

The sample consisted of 30 subjects (13 female). The mean age of the subjects was 41.5 years (SD=13.2 years).

2.7 Test procedure

After instructions about the partial automated system, the NDRT and a short drive for practicing system deactivation, subjects performed the test drive in their assigned test condition. All drivers were told to direct their attention to the video, but on the premise that they must remain ready to intervene at any time.

2.8 Dependent measures

Reaction time until intervention, reaction time of first glance to the road, mean glance duration to video, number of critical events and perceived usefulness of the system's messages were analyzed.

3. Results

Descriptively, the reaction time until intervention was highest in the baseline condition and lowest in the MR condition. However, this comparison does not become statistically significant ($F[2;27]=2.834$; $p=.076$). Subjects in the baseline condition looked up to the scenario descriptively the latest. However, also this difference does not become statistically significant ($F[2;26]=1.266$; $p=.299$).

Mean glance durations to the video revealed a significant effect of the DIL strategy ($F[2;26]=3.986$; $p=.031$) and significant differences between the MR condition ($p=.027$) and the baseline condition as well as between the MR and the DMS condition ($p=.016$) indicating the shortest durations for the MR condition.

A closer look at the gaze durations per subject shows that gaze behavior differs heavily between individuals: The individual monitoring behavior of the drivers probably overlapped the effects of the respective DIL strategies.

- Situation-dependent strategy (MR condition): The driver received a so-called "monitoring request" about an uncertain traffic situation ("Unclear traffic situation"), which might require a possible intervention. It was triggered simultaneously with the appearance of the situation 10 s before reaching the obstacle (also in the distractor scenarios without an obstacle). The message was displayed below the system status for 2 seconds and was combined with an audible advisory tone.
- No strategy (baseline condition): The driver received no warning at all.

In total, 12 collisions/near-collisions with the obstacle occurred, with a clear accumulation in the baseline condition (10 in baseline vs. 1 in DMS vs. 1 in MR condition).

Collisions or near-collisions with the traffic behind occurred eleven times in total, equally distributed over the different test conditions (3 in baseline vs. 5 in DMS vs. 5 in MR condition).

Perceived usefulness of the messages was significantly higher for the MR than for the DMS warnings ($F[1;18]=15.63$; $p=.001$). Descriptively, drivers from the DMS condition felt slightly more disturbed by the messages than the MR condition, without statistical significance ($F[1;18]=1.72$; $p=.206$).

4. Discussion and conclusion

In summary, there were visible differences depending on the driver-in-the-loop strategy in gaze reaction times for the perception of the situation and intervention times for the reaction to the system limit. However, the effects were overlaid by strong individual differences in monitoring behavior during partial automated driving, so that none of these differences reached statistical significance.

The qualitative analysis of single critical events indicated that in the condition without intervention, critical situations with the obstacle occurred more frequently or, conversely, such situations could apparently be successfully avoided by the strategies. The subjective evaluations revealed that MR was perceived as more helpful in preparing for the upcoming system limit.

The results show further that driver's performance of control glances during the interaction with NDRT is a highly individual and automated process which seems very difficult to be influenced by experimentally induced instructions and variations. In general, the experimental induction of distraction (in contradiction to legal aspects and driver's natural behavior) in order to investigate the effectiveness of possible countermeasures must be discussed.

5. Acknowledgments

This work was funded by the Research Association for Automotive Technology (FAT) within a project with the goal to develop a standardized method for the assessment of intervention/takeover performance during L2/L3 automated driving (project report see Wiedemann et al., 2022). Thanks to all members of working group 2 for the fruitful discussions and comments supporting the development of this method. We also thank the reviewers for their valuable time and effort.

References

- Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J., ... & Morgan, J. F. (2015). Human factors evaluation of level 2 and level 3 automated driving concepts (Report No. DOT HS 812 182). Washington, DC: National Highway Traffic Safety Administration.
- Gold, C., Lorenz, L., Damböck, D., & Bengler, K. (2013). Partially automated driving as a fallback level of high automation. 6. Tagung Fahrerassistenzsysteme. Munich, Germany.
- Louw, T., Madigan, R., Carsten, O., & Merat, N. (2017b). Were they in the loop during automated driving? Links between visual attention and crash potential. *Injury Prevention*, 23, 281-286. doi:10.1136/injuryprev-2016-042155
- Louw, T., Markkula, G., Boer, E., Madigan, R., Carsten, O., & Merat, N. (2017a). Coming back into the loop: Drivers' perceptual-motor performance in critical events after automated driving. *Accident Analysis & Prevention*, 108, 9–18. doi:10.1016/j.aap.2017.08.011
- Lu, Z., Zhang, B., Feldhütter, A., Happee, R., Martens, M., and De Winter, J. C. F. (2019). Beyond mere take-over requests: the effects of monitoring requests on driver attention, take-over performance, and acceptance. *Transp. Res. Part F Traffic Psychol. Behav.* 63, 22–37. doi: 10.1016/j.trf.2019.03.018
- Merat, Natasha, Bobbie Seppelt, Tyron Louw, Johan Engström, John D. Lee, Emma Johansson, Charles A. Green, et al. "The 'Out-of-the-Loop' Concept in Automated Driving: Proposed Definition, Measures and Implications." *Cognition, Technology & Work* (September 15, 2018).
- Schömig, N. & Kaussner, A. (2014, July). Evaluation of an Attention Monitor for detection of and feedback on driver's distraction level on different automation levels. Paper presented at the 5th International Conference on Applied Human Factors and Ergonomics (AHFE), Krakow, Poland.
- Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation expectation mismatch: Incorrect prediction despite eyes on threat and hands on wheel. *Human Factors*, 60(8), 1095–1116. doi:10.1177/0018720818788164
- Wiedemann, K., Schömig, N., Wehner, T., Befelein, D. & Neukum A. (2022). Entwicklung einer standardisierten Prüfanordnung zur Bewertung der Übernahmeleistung beim automatisierten Fahren. FAT-Schriftenreihe Band 356. Available at: <https://www.vda.de/de/aktuelles/publikationen/publication/fat-schriftenreihe-356>

European NCAP assessment approach on Occupant Status Monitoring

Adriano Palao^{1*}, Colin Grover^{1,2}, Rikard Fredriksson^{1,3,4}

¹ European New Car Assessment Programme (Euro NCAP), Leuven, Belgium

² Thatcham Research, Berkshire, United Kingdom

³ Swedish Transport Administration, Skövde, Sweden

⁴ Chalmers University of Technology, Gothenburg, Sweden

* adriano_palao@euroncap.com

Abstract: Euro NCAP presents a pragmatic 2-stage method to assess Driver State Monitoring systems, of which direct monitoring is to make a first landing in the five star safety rating scheme from 2023 onwards: First, Euro NCAP reviews a comprehensive dossier provided by the Original Equipment Manufacturer (OEM). Second, the official test laboratory spot-tests some of the system functionalities as required in the protocol.

1. Introduction

Analysis of driver inattention using naturalistic driving studies [1] show that engaging in visually demanding tasks (even for 2-second glances) and driving while drowsy result in higher near crash/crash risk [2]. In addition, increasingly busy global traffic environment requiring attentiveness for safe driving, combined with the trend of vehicles being marketed on their connectivity, and with ever growing screen sizes loaded with potentially distracting features, it becomes necessary for technology to encourage safe and attentive driving.

Acknowledging these facts, and in its constant effort to raise the bar for vehicle safety standards, the European New Car Assessment Programme (Euro NCAP) deems it essential to reward driver monitoring systems that effectively detect impaired driving in the form of driver distraction and drowsiness, as well as initiating various strategies that warn the driver and/or mitigates risks. To that end, direct monitoring will make a soft landing in the 2023 implementation of the “Safety Assist – Safe Driving” protocol, and will be required in order to achieve the full score in the Occupant Status Monitoring (OSM) area. From 2025 onwards, only direct monitoring systems will be rewarded, opening the door to rating more advanced features such as e.g. Cognitive Distraction, Driving Under Influence, more fool-proof seat belt reminders and occupant classification for passive restraint optimization.

The 2023 protocol requires the DSM (Driver Status Monitoring) system to perform when put under a defined range of specific noise variables e.g. driver attributes, as well detecting the status of the driver and including system intervention strategies. As the requirements are difficult to test in a repeatable and reproducible way on a test track, Euro NCAP has put in place a pragmatic approach to assess the performance of DSM systems.

2. Method

The assessment approach consists of two stages: First, Euro NCAP Secretariat will review a dossier provided by the OEM, which summarizes in detail the DSM system performance across the requirements summarized in the protocol (and beyond); Second, an Euro NCAP official test laboratory conducts spot testing on a proving ground where

randomly selected aspects of the DSM system are checked to confirm functionality.

2.1 DSM Dossier Guidance

Euro NCAP has elaborated a Technical Bulletin (TB) supplementary to the protocol, which provides provisions and guidance [3] on the format, minimum contents and structure of the DSM dossier. Some of the minimum required sections and provisions for the OEM to include in the dossier are described below.

2.1.1 System Overview

Summary of the main system functionalities, compliance of the minimum system requirements, sensors involved in the system, their role and relevant specifications, and details explaining the constituent elements of the different system warnings;

2.1.1 Noise Variables

Containing compelling evidence that the system can monitor a population constituted of different types of drivers, with a range of facial occlusions and driver behaviours. Depending on the complexity of the noise variables, the requirement vary between ‘Must’, ‘Inform driver if degraded’, and ‘Information only’;

2.1.1 Detection of driver state

Supporting evidence demonstrating that the system can effectively classify the driver state in the minimum required categories:

- Distraction: further classification of distraction includes ‘long distraction’, ‘short distraction’, and ‘phone usage’. As distraction is heavily linked to gaze location, the OEM is required to specify in the dossier a drawing the delimited gaze areas/regions which the system takes into account to assess distraction.
- Fatigue: further classification of fatigue includes ‘drowsiness’, ‘microsleep’ and ‘sleep’. Euro NCAP gives freedom to the OEM to include in the dossier other methods to assess fatigue other than the ones specified in the protocol.

- Unresponsive driver: details of how the driver status is deemed unresponsive (or sudden sickness) by the system

2.1.1 Vehicle response requirements

Including details on how the sensitivity of ADAS is increased (e.g. Forward Collision warning – FCW; Lane Departure Warning – LDW) when driver is deemed distracted, fatigued or unresponsive. The OEM is free to stick to the protocol requirements or justify other vehicle response methods;

2.2 DSM Spot Testing Guidance

Complementary to the information provided by the OEM in a dossier, the spot testing is the second stage in the assessment of the DSM performance. Euro NCAP has consolidated a comprehensive guideline [4] with the necessary provisions on how the spot testing is to be conducted across official test laboratories.

2.2.1 Test provisions

The test is to be conducted under defined conditions so as to maximize repetitiveness across test laboratories (e.g. uniform surface with consistent slope, at daylight without direct glare or strong light transitions, avoiding strong precipitation).

The vehicle under test (VUT) is to be instrumented with a relatively simple measuring equipment, recording at a defined sample rate (>25Hz): the VUT speed, driver’s gaze location and DSM warning(s). Time variables are defined to ensure consistency, and are to be used later for analysis purposes (Table 1). Furthermore, prior to the test, the timing of FCW and LDW are to be checked at their minimum operational speed without signs of driver inattentiveness, so that the sensitivity increase can be later assessed. It is also important to ensure that previous system learnings on driver drowsiness are reset.

Table 1 Variables to be measured during Spot Testing

| Variable | Description |
|----------------------|--|
| T | Time |
| T ₀ | Time when manoeuvre starts |
| T _{away} | Time of first eye movement away from forward road view |
| T _{gaze} | Time of first glance on gaze location |
| T _{dist} | Time where distracted warning activates |
| T _{warn} | Time of first instance of audio/visual warning |
| T _{fat} | Time where fatigue warning activates |
| T _{FCW} | Time where FCW activates with attentive driver |
| T _{FCWdist} | Time where FCW activates with distracted driver |
| T _{LDW} | Time where LDW activates with attentive driver |
| T _{LDWdist} | Time where LDW activates with distracted driver |

2.2.2 Test execution

The test laboratory in charge of the assessment will randomly pick a test subject (a qualified driver from their staff) whose variables and ranges are within the protocol specifications. The driver will then adjust the seat in the preferred position, and proceed with the test after the vehicle preparation.

Euro NCAP secretariat will ask the test laboratory to spot test a number of distraction, fatigue and unresponsive driver areas of the DSM system, which performance has been claimed in the dossier by the OEM. While the vehicle is in motion at a defined constant speed deemed adequate for the test, the driver shall keep a defined head and body posture while looking to the road ahead, until the manoeuvre begins.

For distraction scenarios, the driver will proceed with moving the head, eye gaze or body posture (depending on the scenario) towards the target area (e.g. glovebox, side mirror, rear passenger seat, etc), and hold the position for a defined time as required in the protocol. An extra time of +1 second is added to the required time, so as to ensure that the system reaction is captured during the assessment.

For the assessment of Fatigue and unresponsive driver, Euro NCAP reserves the right to investigate it in practice, although it should rely on the evidence reflected in the dossier. For microsleep, sleep and unresponsive driver scenarios, the metrics by default for assessment are eye closure timing and eventually head nodding forwards; however, a different OEM strategy is allowed for as long as it is justified;

Finally, for each of the areas where the system was functional, the scenario will eventually have to be repeated with different occlusions (cap, hat, sunglasses and facemask).

3. Results

The DSM Dossier Guidance TB elaborated by Euro NCAP aims to ease the reviewing process by standardizing the document across different DSM systems, while granting enough flexibility for the OEM to include additional detail deemed necessary to further illustrate the system constituent components, performance and functionalities (e.g. schematics, diagrams, videos). Euro NCAP also gives the OEM room to use other methods to accomplish the different DSM system requirements, for as long as a details justifying the safety benefits are included in the dossier. Furthermore, the dossier will provide Euro NCAP test laboratories an overview of the particularities of the system to be spot-tested.

By encouraging the OEM to elaborate an in-depth analysis of their system in the dossier, Euro NCAP aims to get valuable insights for more holistic learnings, leveraging the development of future protocols.

The DSM Spot Testing Guidance TB makes sure that repeatability is maintained across test laboratories and systems.

4. Discussion

The presented method is subjected to further refinements as gaps, inconsistencies or feasibility issues are found alongside the 2023 test campaign. Euro NCAP will constantly keep track of those and liaise with test laboratories

and the involved industry members so as to enhance the method. Furthermore, it is expected that the method enables deeper understanding of system capabilities and of possible improvements in requirements for future assessment upgrades.

5. Conclusions

The presented method provides a pragmatic approach to understand the system functionalities and assess the performance of those when put under the minimum requirements to score points in the Euro NCAP DSM protocol.

6. Acknowledgments

Thanks to Sjef van Montfort and James Ellway for the guidance and support during the creation of the TB's.

References

- [1] Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J., "The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car naturalistic driving study data. Report No. DOT HS 810 594," National Highway Traffic Safety Administration, Washington, D.C., 2006.
- [2] Victor T., Dozza M., Bärghman J., Boda, C.N., Engström J., Markkula G., "Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk (SHRP 2)," Transportation Research Board, Gothenburg, Sweden, 2014.
- [3] Palao, A., "TB 036 - DSM Dossier Guidance," Euro NCAP, Leuven, Belgium, 2022.
- [4] Grover, C., "TB 039 - DSM Spot Testing Guidance," Euro NCAP, Leuven, Belgium, 2022.

How does trust in vehicle automation affect the handover process? A Wizard-of-Oz study on public roads

Thomas Streubel^{1*}, Fredrick Ekman², Mikael Johansson²

¹ Division Vehicle Safety, Chalmers University of Technology, SE-412 96, Gothenburg, Sweden

Current affiliation: Volvo Car Cooperation AB, Safety Center, SE-405 31, Gothenburg, Sweden

² Division Design & Human Factors, Chalmers University of Technology, SE-412 96, Gothenburg, Sweden

*thomas.streubel@gmail.com

Abstract: Trust in vehicle automation has been established as an important prerequisite to user acceptance, which in turn leads to increased usage. This enables the expected benefits of automated driving functions (ADFs), especially in terms of traffic safety. This study explores driver behavior after activating automated driving (observed and self-reported) as well as self-reported trust features after automated driving in real traffic. A potential context is investigated to what extent trust influences the speed in which the driver disengages from the driving task after activating automation. The automation is realized with a Wizard-of-Oz vehicle controlled by a wizard driver in the back seat. A total of 30 participants were driving on the city highway of Gothenburg in Sweden. The results show high correlation of trust features and overall elevated levels of trust after the drive. Self-reported activities with ADFs also indicate an expected shift in behavior towards non-driving related tasks such as texting, browsing and in general using the smartphone extensively. The observed handover process shows a fast activation and disengagement from the driving task. Most drivers freely engaged in other activities as there were no specifications on what to do while in automated mode. The drivers with higher trust levels were more likely to engage earlier in non-driving related tasks (in respect to time after activation). The results also show that high trust in automation co-occur with a fast disengagement from the driving task. This has potential implications on the function development and the expectation of a fallback-ready driver. While drivers stated they like to monitor the system, most started to engage in secondary activities with little to no monitoring. This may be an effect of seeing the system perform for the first time (halo effect).

1. Introduction

As the progression from partial to fully automated vehicles (AVs) accelerates, the driver's role may eventually change from that of active operator to that of passenger. This is expected to lead to increased traffic safety, traffic flow, comfort (for user) and insure mobility for all (old and impaired users) (Kyriakidis et al., 2017). However, for a successful change of the driver's role, it is important for the user to trust the AV, since trust is believed to be a prerequisite for acceptance (Ghazizadeh et al., 2012; Molnar et al., 2018; Zhang et al., 2020) and adoption of AVs (Choi & Ji, 2015).

An important attribute to establish trust is what the automation does in order to reach the user's goals, and includes factors such as reliability, predictability and ability (J. D. Lee & See, 2004). Reliability concerns how consistent the automation performs, and predictability is defined as how well the automation performs according to the user's expectations (Hoff & Bashir, 2015). It has been shown that the design of an AV driving behavior affects users perception of how predictable an AV is, which in turn affects trust (Ekman et al., 2019).

Questionnaires (Jian et al., 2000; Li et al., 2019) and interviews (J. Lee et al., 2016) are often used to measure users trust. Other researchers have explored the potential of using behavioral indicators to measure users' trust. Previous studies used indicators such as gaze behavior (Hergeth et al., 2016; Walker et al., 2018), head-, hand- and foot position (Wright

et al., 2016). However, the results from the studies are inconclusive regarding the correlation between objective and subjective trust measurements. Furthermore, the studies also contain shortcomings in their experimental setup, i.e. simulator studies that lack perception of risk which is a fundamental aspect for trust to exist (J. D. Lee & See, 2004).

Thus, the aim of this work is to present the results from an experiment regarding how users behavior relates to trust during handover situations (transition from manual to automated driving). These results include both questionnaire ratings as well as behavioral indicators such as eye-gaze, head-, hand- and foot position during the transition as indicators of trust in AVs in a naturalistic driving environment i.e., in real traffic.

2. Method

The study was conducted with a Wizard-of-Oz (WoZ) vehicle on public roads. The participants were under the impression that they were testing an automated driving function (ADF). During the drive, the control switched twelve times between the driver and the function in designated locations. The drivers could activate automation (handover) and were required to retake control (takeover).

2.1 Participants

There were 30 participants (10 female) available for analysis in this study. All were employees of Volvo Car Cooperation for legal issues. The age ranged from 23 to 64 ($M = 39.2$, $SD = 10.5$). The driving experience was above 2

years for all participants with a median yearly mileage of 15 to 20 thousand kilometers.

2.2 Test vehicle

The test vehicle was a modified Volvo XC90 with double controls. The "wizard" controls the car when in automated driving mode sitting in the center of the rear bench seat. For this purpose, there is a steering wheel as well as pedals connected to the vehicle's CAN bus (drive-by-wire). The wizard driver can also utilize the built-in Adaptive Cruise Control (ACC). The controls are obscured with a cover to be not visible by the driver. In addition, there are three cameras placed in the cabin to capture the driver's behavior.



Figure 1. Route on the city highway ring in Gothenburg with AD stretches marked in blue

2.3 Route

The test drive was performed on the outer ring road around Gothenburg, Sweden (see dashed line in Figure 1). One round is approximately 30 kilometers long. The road is a city highway with 2-3 lanes in each direction separated by a median barrier. The posted speed limit is 70 to 80 km/h. The traffic is moderate during the day with density spikes during common rush hours.

2.4 Procedure

All drivers had a briefing before the drive to get familiar with the handling of the vehicle. The wizard in the back was introduced as a backup for a system failure. The drivers were allowed to engage in non-driving related tasks (NDRTs). There was also a tablet mounted on the center stack to be used freely during automation. The drive was two rounds clockwise which took about one hour. Handovers were initiated by the test leader and indicated in the Driver Information Module (DIM). The drivers were asked to use the ADF whenever available. After the drive, there was a questionnaire and a short debriefing to determine if the drivers were unaware of the purpose of the wizard.

2.5 Design

Overall, there were 6 phases in automated driving mode, phases 1 and 4 are about 1 minute short (test phases). The handover process is compared across all phases by video annotation of behavioral attributes, such as *first glance on the*

DIM, activation of automation, hands off wheel, feet resting position, start secondary task.

The questionnaire statements related to trust were on a 5-point Likert scale. The overall trust is evaluated at the end of the drive and correlations with different trust features were assessed using the Spearman coefficient.

2.6 Data collection and validation

The driver video data was annotated for 30 seconds after the system became available based on the attributes mentioned above. Three annotators were used to validate the subjective data collection by cross comparison of inter-rater reliability.

3. Results

3.1 Features of trust

The drivers assessed the system predominantly positively. *Trust, reliability and intention to rely* are high with more than 50% strongly agreeing. *Ability and predictability* are slightly lower, with both having a neutral spectrum of about 15%. The *desire to monitor*, in which nearly half of the participants agree or strongly agree, has also a neutral spectrum of about 30%. The ratios are displayed in Figure 2.

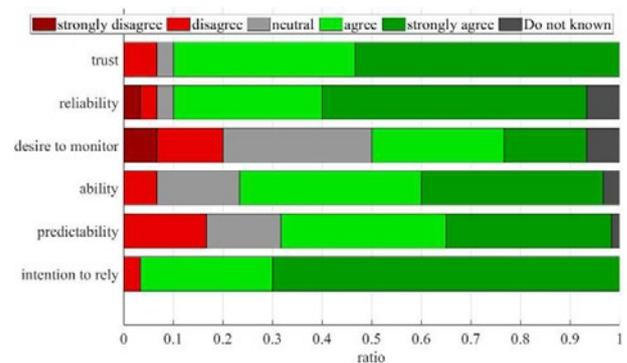


Figure 2. Results of the trust related items in questionnaire

All features showed weak to moderate correlations with trust (see Table 1). While most trust related features also show correlation with each other, *desire to monitor* stands out with no other relation other than a weak correlation with trust.

Table 1. Spearman correlation values between the features extracted from the questionnaire

| feature | Intention to rely | Predictability | ability | Desire to monitor | Reliability |
|-------------------|-------------------|----------------|---------|-------------------|-------------|
| Predictability | 0.14 | | | | |
| Ability | 0.52 | 0.24 | | | |
| Desire to monitor | -0.1 | -0.03 | -0.03 | | |
| Reliability | 0.39 | 0.43 | 0.46 | -0.06 | |
| Trust | 0.45 | 0.27 | 0.31 | 0.27 | 0.53 |

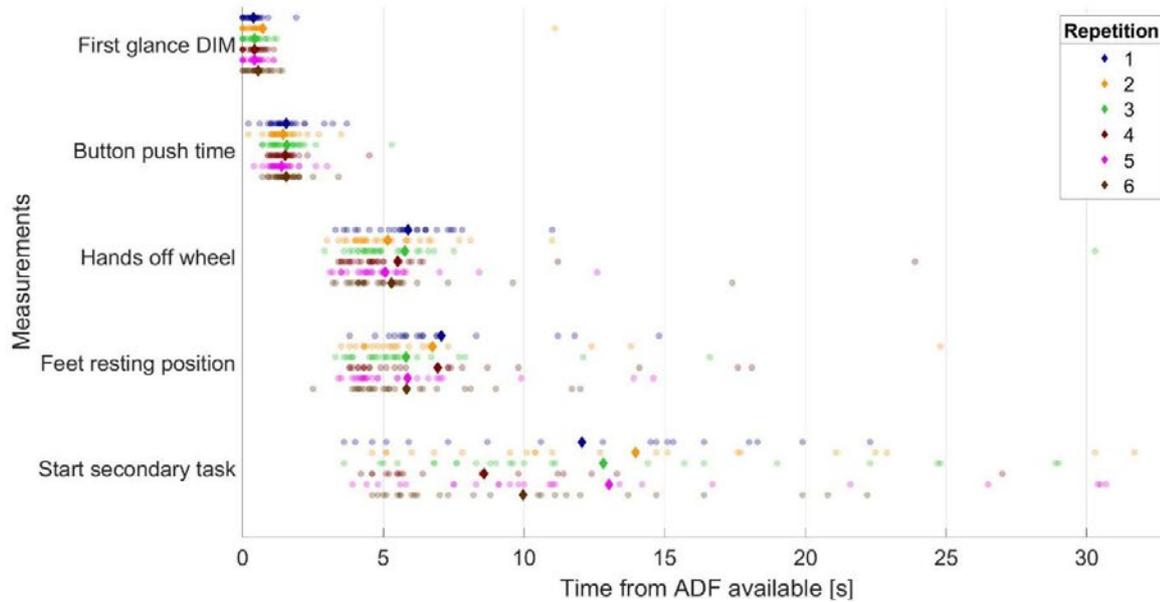


Figure 3. Timing of the response measures taken after automated driving function is available (bold markers indicate means)

3.2 Differences in handover behavior

The automated driving function was enabled by the test leader in the passenger seat, and an acoustic and visual signal was sent to the driver via the DIM. Then the handover process was observed, marking the timings of the behavioral features of interest (see Figure 3).

The *first glance DIM* occurred within 2 seconds for the 95-percentile of drivers (in all repetitions). This was followed by an immediate system activation indicated by *button push time*. After only 5 to 6 seconds on average drivers took *hands off wheel* and most drivers took the feet of the pedals in a *feet resting position*. Most drivers *start secondary tasks* within the 30 second observation span. The timing is more spread over the different repetitions of automation phases. Comparisons with the self-reported trust showed a correlation of trust with the onset of secondary tasks.

4. Discussion

The participants were highly trusting in the automation after experiencing it in real traffic. However, the *desire to monitor* was rather high even among trusting participants. This might be an effect of wanting to see the system operating for the first time (halo effect) and potentially will degrade over time.

The actual behavior observation does not show high attentiveness after handover. Only few drivers switched into the role of supervising the automation. Although, supervision is not required in conditional automation (level 3) or higher, it is deemed to be useful by SAE level definition (Society of Automotive Engineers (SAE), 2021). In contrast to previous simulator studies, where in vehicle tasks are introduced by design, the drivers in our real traffic environment showed naturally shifts towards secondary activities while in automation mode (Jamson et al., 2013). Also, the surrounding

traffic density seemed to have no influence on the level of attentiveness away from the driving task.

Especially, the activation time is fast which indicates eagerness to try the system on one side but also is a result of the instruction to activate as soon as possible. Interestingly, it takes only 6-8 seconds on average (after automation available) until the controls are given up completely shown by the driver leaving the hands off the wheel and feet in a resting position (away from the pedals). There was also no effect of repetition in these actions as the timing does not change significantly between different automation phases.

That the self-reported trust correlates with the onset of starting a secondary task indicates, that after trust is established, drivers will quickly disengage from the driving task and not monitor further nor give attention to the driving environment. This is a key factor when it comes to ADF design and the expectation of having a fallback-ready user in case of deactivation.

5. Conclusions

The highly trusting participants in this study handed over the control to the vehicle in a fast and smooth way, leaving the controls shortly after. The disengagement from the driving tasks often leads to an attention switch towards secondary activities. It cannot be expected that a driver trusting in an automated driving function will frequently monitor the system and have some level of situational awareness in relation to the driving task. This should be considered in the design of such functions.

6. Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 723051. The sole responsibility of this publication lies with the author(s). The author(s) would like to thank all partners within L3Pilot for their cooperation and valuable contribution.

References

- Choi, J. K., & Ji, Y. G. (2015). Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *International Journal of Human-Computer Interaction*, 31(10), 692–702. <https://doi.org/10.1080/10447318.2015.1070549>
- Ekman, F., Johansson, M., Bligård, L.-O., Karlsson, M., & Strömberg, H. (2019). Exploring automated vehicle driving styles as a source of trust information. *Transportation Research Part F: Traffic Psychology and Behaviour*, 65, 268–279. <https://doi.org/10.1016/j.trf.2019.07.026>
- Ghazizadeh, M., Lee, J. D., & Boyle, L. N. (2012). Extending the Technology Acceptance Model to assess automation. *Cognition Technology and Work*, 14(1), 39–49. <https://doi.org/10.1007/s10111-011-0194-3>
- Hergeth, S., Lorenz, L., Vilimek, R., & Krems, J. (2016). Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving. *Human Factors*, 3, 509–519. <https://doi.org/10.1177/0018720815625744>
- Hoff, K., & Bashir, M. (2015). Trust in automation: integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3). <https://doi.org/10.1177/0018720814547570>
- Jamson, A. H., Merat, N., Carsten, O. M. J., & Lai, F. C. H. (2013). Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. *Transportation Research Part C: Emerging Technologies*, 30, 116–125. <https://doi.org/10.1016/j.trc.2013.02.008>
- Jian, J.-Y., Bisantz, A., & Drury, C. (2000). No Title. *International Journal of Cognitive Ergonomics*, 4(1), 53–71. https://doi.org/10.1207/S15327566IJCE0401_04
- Kyriakidis, M., de Winter, J. C. F., Stanton, N., Bellet, T., van Arem, B., Brookhuis, K., Martens, M. H., Bengler, K., Andersson, J., Merat, N., Reed, N., Flament, M., Hagenzieker, M., & Happee, R. (2017). A human factors perspective on automated driving. *Theoretical Issues in Ergonomics Science*, 0(0), 1–27. <https://doi.org/10.1080/1463922X.2017.1293187>
- Lee, J. D., & See, K. a. (2004). Trust in automation: designing for appropriate reliance. *Human Factors*, 46(1), 50–80. <https://doi.org/10.1518/hfes.46.1.50.30392>
- Lee, J., Kim, N., Imm, C., Kim, B., Yi, K., & Kim, J. (2016). A Question of Trust: An Ethnographic Study of Automated Cars on Real Roads. *8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotiveui 2016)*, 201–208. <https://doi.org/10.1145/3003715.3005405>
- Li, M., Holthausen, B., Stuck, R., & Walker, B. (2019). No Risk No Trust: Investigating Perceived Risk in Highly Automated Driving. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 177–185. <https://doi.org/10.1145/3342197.3344525>
- Molnar, L., Ryan, L., Pradhan, A., Eby, D., St Louis, R., & Zakrajsek, J. (2018). Understanding trust and acceptance of automated vehicles: An exploratory simulator study of transfer of control between automated and manual driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 58, 319–328. <https://doi.org/10.1016/j.trf.2018.06.004>
- Society of Automotive Engineers (SAE). (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*.
- Walker, F., Verwey, W., & Martens, M. (2018). Gaze Behaviour as a Measure of Trust in Automated Vehicles. *Proceedings of the 6th Humanist Conference, The Hague, Netherlands, 13-14 June 2018*.
- Wright, T., Horrey, W., Lesch, M., & Rahman, M. (2016). Drivers' trust in an autonomous system: Exploring a covert video-based measure of trust. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1), 1334–1338. <https://doi.org/10.1177/1541931213601308>
- Zhang, T., Tao, D., Qu, X., Zhang, X., Zeng, J., Zhu, H., & Zhy, H. (2020). Automated vehicle acceptance in China: Social influence and initial trust are key determinants. *Transportation Research Part C: Emerging Technologies*, 112, 220–233. <https://doi.org/10.1016/j.trc.2020.01.027>

Effect of engagement with a Trivia game on driver's sleepiness and behavioural adaptation in a partially automated vehicle.

Dan Boguslavsky^{1*}, Hadas Chassidim², Jiaqi¹ L, Avinoam Borowsky¹

¹ Industrial Engineering and Management, Ben-Gurion University of the Negev, Ben-Gurion Ave. 1, Be'er-Sheva, Israel

² Shamoon College of Engineering, 56 Bialik St., Be'er Sheva, Israel

*danbogu@post.bgu.ac.il

Abstract – The objective of this study was to use a driving simulator to examine whether the utilization of an activating non-driving related task (NDRT) as a countermeasure can mitigate the effects of passive fatigue development and hazard perception (HP) deterioration under partially automated driving (PAD). Twenty-four participants were randomly assigned to one of two experimental conditions: (1) driving under PAD conditions without an activating NDRT or (2) driving under PAD conditions in the presence of an activating NDRT. The activating NDRT was a multimodal version of a Trivia game (auditory and visual). Each condition included two driving sessions one week apart. Each driving session included four unmaterialized hazardous situations (8 total) counterbalanced between participants. Participants were connected to an eye-tracking system and ECG throughout their drives. In addition, they were asked to rate their subjective level of sleepiness (KSS) four times during each drive. Initial findings show that the progression of subjective sleepiness is slower when drivers can engage with an activating NDRT than when they are not. This effect dissipates as drivers gain experience with the system.

Keywords: Distraction, Partially Automated Driving (PAD), Sleepiness, Countermeasures, Non-Driving Related Task (NDRT).

1. Introduction

The increased prevalence of automation functions in partially automated vehicles (PAV) relieves drivers from vehicle control tasks but not from their role as supervisors of the automated system and the driving task [1]. This supervision requires continuous and passive monitoring, which is monotonous and tedious, resulting in passive fatigue, decreased vigilance, and even compromising the driver's ability to react in a critical event [2,3].

This work was aimed to evaluate whether engagement with an activating non-driving related task (NDRT) can suspend passive fatigue development and at which costs (e.g., increased workload, distraction).

2. Method

2.1 Participants

Twenty-four participants, 12 females (mean age=25.25, SD=2.83; mean driving experience=7.25, SD=2.7) and 12 males (mean age=27.3, SD=7.03; mean driving experience=9.63, SD=7.26). All participants were students at Ben-Gurion University of the Negev (BGU) and had more than five years of driving experience. In addition, participants went through visual acuity (Snellen Chart), and contrast sensitivity (FACT; Ginsburg, 2003) tests to assure they had a normal, or corrected-to-normal, vision (6/9 or better and normal contrast sensitivity function). Participants also declared they do not suffer from any cardiological problems, light sensitivity, or a tendency to headaches and nausea. The participants had received 140 NIS after completing the experiment. The BGU IRB ethically approved the study.

2.2 Apparatus

2.2.1 Driving Simulator

An RTI high fidelity driving simulator (Realtime Technologies, Inc.) was used for the study. The driving simulator consists of an engineless Cadillac-STs sedan and a 7m diameter curved screen (2.4m X 6.1m), creating a visual angle of 165 degrees of the virtual world, located at about one meter in front of the vehicle. Three laser projectors displayed the virtual world on the curved screen, and a designated software (Wrapalizer, Inc.) did the edge blending. A rear projector and a screen at the back of the simulator presented the virtual environment through the in-vehicle rear-view mirror. In addition, each physical side mirror included a 7" LCD showing the respective views of the virtual environment.



Figure 1 - RTI high fidelity driving simulator

2.2.2 Human Machine Interface (HMI)

Two in-vehicle displays were connected to a PC and were used to display the HMI. One was located behind the steering wheel and was used as a digital instrument panel dashboard to display information and notify the drivers about hazard scenarios, fatigue, and automation level. The other was located in the central console and was used as an infotainment screen. The infotainment screen displayed the online KSS questionnaire, visual and vocal alerts about on-road hazards and fatigue status, and the Trivia game interface.

2.2.3 Eye tracker and Electrocardiogram (ECG)

Participants' eye movements were recorded with TOBII head-mounted pro glasses 2.0 at 50Hz. In addition, participants' heart rate was recorded with A BioPac ECG system (MP150) at 2000Hz.

2.2.4 Driving Environment and Scenarios

Each drive session included a trip of approximately 40 minutes, usually on a straight urban or highway road with sparse traffic. During each driving session, four unmaterialized hazard scenarios occurred along the road. Different combinations of the scenarios were assigned randomly between the participants. Each scenario presented a latent cue that could be spotted from a distance.

2.2.5 Training Drive

Before the experimental drive, participants underwent a 5-minutes driving training session and received an explanation of the environment, devices of the simulator, and a demo of the HMI.

2.2.6 Questionnaires

Participants were asked to complete three sets of questionnaires: (Set1) included (1.1) demographics, (1.2) previous experience and familiarity with automation, and (1.3) adoption and trust. (Set2) was administered after training and included knowledge verification regarding the simulated L2 driver functionalities. (3) Post-drive set included (3.1) workload – NASA TLX, (3.2) usability of the Mediator system, (3.3) Knowledge verification regarding Mediator's HMI functionality, and (3.4) adoption and trust (same as 1.3).

2.2.7 Secondary Task

The secondary task interface consists of a multiple-choice questions trivia game. Each time a participant accepts an invitation to play trivia, a batch of 11 questions follows. The content of the question and the possible answers were read out loud with speakers inside the car. The driver chose an option by pressing the respective button on the central touchscreen.

2.3 Experimental Design

The experiment was a mixed 2-by-2 factorial design. Participants were randomly assigned to one out of two experimental conditions: (1) L2 driving with an NDRT (2) L2 driving without an NDRT. This was a between-subjects independent variable. Gender was balanced within each

experimental condition. Both conditions interacted with the HMI system and were asked to evaluate their subjective sleepiness level, received visual and vocal alerts of hazards ahead, automation status, and alerts in case of fatigue. Participants from the first experimental condition were offered three times to engage with a Trivia game as an NDRT throughout the drive (see Figures 2&3). In contrast, the HMI design was adapted to the driver's sleepiness level based on the KSS scores. When a driver was detected as sleepy for the first time, an "eye" icon notification was displayed. If the driver kept reporting high KSS scores in the following KSS instance, then a "coffee cup" icon appeared on the central display, asking the driver to stop for a rest.

Participants drove two driving sessions in the simulator one week apart in each experimental condition. The long-term HMI effect was a within-subject independent variable.

2.4 Procedure

When arriving at the lab, participants declared that they were well-rested, asked to report a KSS score, and performed two vision tests.

Qualified participants received written instructions regarding the simulator automation capabilities, the simulated environment, the HMI infotainment system, and the measurement apparatus. Participants were told that they should drive as they would in similar real-world situations, and they should constantly monitor the automated system and the driving task. Participants were told that the driving task is under their sole responsibility.

Then, participants were connected to an ECG, entered the vehicle, and were asked to sit and relax for 5 minutes while reading a magazine. Meanwhile, a baseline ECG including R-R interval measurement was performed. Next, they wore eye-tracking glasses for calibration.

The participants drove a 5-minutes training session to familiarize themselves with the simulator and experience its behaviour and were briefly introduced to the HMI. Then, the participants were asked to complete a short questionnaire regarding the simulator operation and HMI. Immediately after, the participants started the 40 minutes experiment drive depending on their experimental condition. After the drive, the participants were asked to fill out the post-drive questionnaire.

A week later, the participants returned to the lab approximately at the same hour as the first session and underwent the same procedure as in the first session. After completion, the participants were thanked for their participation and received monetary compensation.

3. Results

Only the mean KSS scores of both experimental conditions were analysed as initial results due to the submission deadline. In general, both experimental conditions demonstrated an increased trend of subjective sleepiness along the drive. However, during the first drive, the first experimental condition (with Trivia) demonstrated a lower rate of subjective sleepiness development than the second condition (Figure 2). The same effect is observed in the second driving session until the third KSS instance. Then, the subjective sleepiness score of the control group (second condition) reaches a ceiling, and the score of the experimental group keeps its trend (Figure 3).

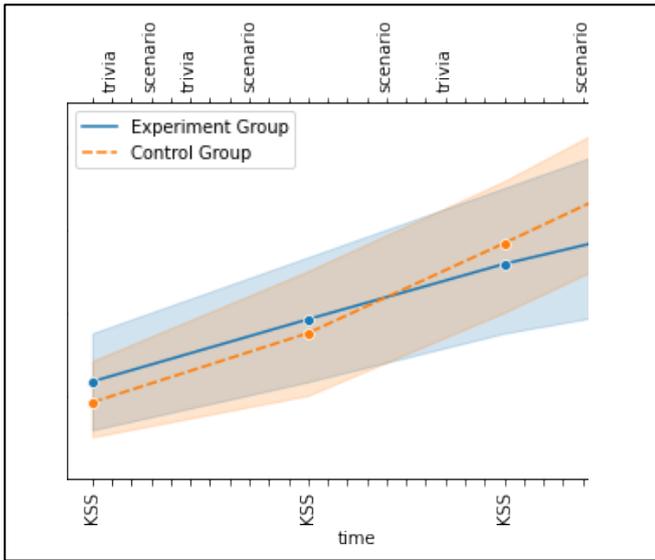


Figure 2 - mean KSS score for the first drive

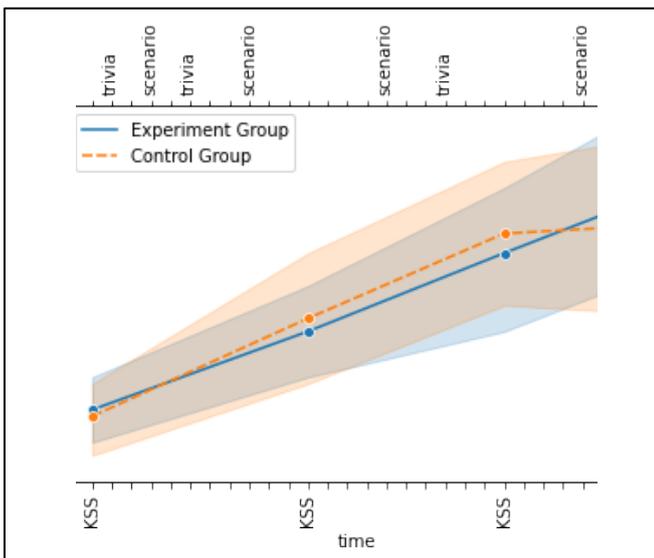


Figure 3 - mean KSS score for the second drive

4. Conclusions

Our findings extend previous findings by showing that using Trivia as a countermeasure for passive fatigue inhibited sleepiness progression, but only for a limited period. After a certain period, there was a 'jump' in sleepiness ratings; thus, the NDRT effect seemed to have dissipated towards the end of the drive. In addition, the KSS scores analysis show that drivers adapt to the Trivia, and its potential to inhibit fatigue progress is reduced between drives.

5. Acknowledgments

This study was performed within the MEDIATOR project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 814735. The authors are solely responsible for the study's contents, which do not necessarily represent the official views of the Mediator EU project.

References

Fisher, D. L., Lohrenz, M., Moore, D., Nadler, E. D., & Pollard, J. K. (2016). Humans and intelligent vehicles: The hope, the help, and the harm. *IEEE Transactions on Intelligent Vehicles*, 1(1). <https://doi.org/10.1109/TIV.2016.2555626>

Ginsburg, A.P., 2003. Contrast sensitivity and functional vision. *Int. Ophthalmol. Clin.* 43 (2), 5–15.

Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving. *Procedia Manufacturing*, 3. <https://doi.org/10.1016/j.promfg.2015.07.499>

Merat, N., Seppelt, B., Louw, T., Engström, J., Lee, J. D., Johansson, E., Green, C. A., Katazaki, S., Monk, C., Itoh, M., McGehee, D., Sunda, T., Unoura, K., Victor, T., Schieben, A., & Keinath, A. (2019). The “Out-of-the-Loop” concept in automated driving: proposed definition, measures, and implications. *Cognition, Technology and Work*, 21(1). <https://doi.org/10.1007/s10111-018-0525-8>

Do drivers engage in NDRTs when travelling in an AV in real-world motorway environments?

Xian Liu^{1,2,3,4*}, Ruth Madigan⁴, Ehsan Sandraei⁴, Yee Mun Lee⁴, Natasha Merat⁴

¹ Jiangsu Key Laboratory of Urban ITS, Southeast University, Nanjing, 211189, China

² Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, Nanjing, 211189, China

³ School of Transportation, Southeast University, Nanjing, 211189, China

⁴ Institute for Transport Studies, University of Leeds, Leeds, United Kingdom

*liuxian@seu.edu.cn

Abstract: Previous simulator and real-world studies with SAE Level 2 automated vehicles (AVs) have shown that, when compared to manual driving, drivers are more inattentive when using partially automated driving systems, reflected by less glances towards the road, and more focus on non-driving related activities. Manual driving studies also suggest that drivers are more likely to engage in non-driving related tasks (NDRTs) during slow moving or stationary traffic conditions. The aim of the current study was to understand if these attention patterns also happen in real-world SAE Level 3 AVs, when supervised by a safety driver. In particular, the study's aim was to understand whether NDRTs engagement is impacted by the driving environment. 46 video clips of drivers interacting with NDRTs during L3 motorway driving were analysed, with speed of travel used as a proxy measure for the volume of surrounding traffic. The number and duration of glances towards the forward roadway were used as proxy measures for engagement in NDRT. Results will be presented by comparing the number and duration of glances for three different driving speeds, to understand the implications of different HMI and automation conditions on how drivers self-regulate their engagement in NDRTs when driving in a real-world SAE Level 3 AV.

1. Introduction

Vehicles which allow some automated functionality, such as lateral and longitudinal support, are now available on the market (e.g., autosteer, lane keeping system, adaptive cruise control). However, these features are still limited, for example, due to a sudden change in weather or traffic conditions (based on the vehicle's Operational Design Domain or ODD). Therefore, according to the Society for Automotive Engineers (SAE, 2021), drivers must "constantly supervise these support features" for Level 2 functions, or drive "when the feature requests" for L3 functionalities. However, L2 and L3 driving simulator studies have shown that, compared to manual driving, drivers' gaze is less focused towards the road centre (which gravely affects supervision- and safety, Goncalves et al., 2020; Louw et al., 2019; Zeeb et al., 2016). Similar results have been reported in real-world L2 automated driving studies, with less glances towards the road, and more focus on non-driving related activities (e.g., engaging with mobile phones), when compared to manual driving (Morando et al., 2021; Noble et al., 2021). However, due to their absence from the road, there is currently little knowledge of how drivers will behave in more advanced, L3 AVs, where they are legally allowed to engage in NDRTs, but must be ready to drive, when requested. If your paper does not meet all of the requirements, your paper will be unsubmitted. It is at the discretion of the Organisation Committee to decide if a submitted contribution that is "unsubmitted" is returned with a request for revision to address identified issues, or if it is simply rejected.

Previous naturalistic studies on manual driving have shown that the driving context and environment are critical factors for influencing drivers' engagement in NDRTs. For

example, drivers are more likely to initiate phone calls, or engage in visual manual phone activities in stationary traffic, compared to higher speed driving conditions (Christoph et al., 2019; Funkhouser et al., 2012; Tivesten et al., 2015). However, most of these studies are based on data from manual driving conditions, where the driving demand is significantly different, compared to automated driving. It can be argued that with an increase in the levels of automation, there is even less demand on the driver, further encouraging their chances to engage in NDRTs. How such levels of automation affect driver engagement in NDRTs, and how the contextual environment, such as road type and traffic conditions affect this engagement is also not known.

The aim of the present study was to investigate the pattern of drivers' engagement in NDRTs, when travelling in an L3 automated test vehicle, on a European highway, with different levels of traffic. The following research questions were investigated:

(1) What is the attention pattern (number and duration of glances towards the forward roadway) when drivers engage in NDRTs in a real-world L3 AV?

(2) How does the driving environment (based on speed of travel) influence this NDRT engagement?

2. Method

2.1 Participants

Thirty-one non-professional drivers (25 females, 6 males) aged 25-70 years ($M_{age} = 40.42$ years, $SD_{age} = 12.22$) took part in this study. They received between €200- and €250-worth of shopping vouchers for this study.

2.2 Automated vehicle and route

The study was conducted between January 2020 and March 2021, on a 95km long motorway section outside a busy European city. The road environment included both busy sections of traffic, and free motorway driving. All experiments took place in clear daytime weather, with no heavy rain or snow, and the drive lasted between 1 and 1.5 hours.

The drive started in manual mode, until the vehicle reached the main motorway route, and the automated driving function (ADF) was available. The SAE Level 3 was capable of driving in its own lane at the designated speed limit, performing overtaking manoeuvres, and changing lanes at speeds of up to 110 kph. For safety reasons, a manually driven vehicle followed the AV.

Participants' actions were recorded using 3 in-vehicle cameras: one camera positioned on the dashboard to capture their posture and head/facial movements, one camera positioned beside their right shoulder to capture hand position and dashboard information, and another to capture the position of their feet on the pedals. Data from these cameras was linked to the timestamps of the vehicle CAN bus data, which provided travelling speed, acceleration, and vehicle positioning. External cameras also recorded the external environment, but, for GDPR reasons, the quality of images was not suitable for use in this study.

2.3 Experimental procedure & design

This study consisted of three automated drives, on the same route, separated by a few weeks. Additional safety measures such as facemasks were included due to the Covid-19 pandemic restrictions.

Before the first experimental drive, all participants were informed about the experiment, provided their informed consent, and completed a pre-experimental questionnaire. On arrival, the participants were given another briefing on the experimental procedure, were familiarised with the AV's driving functions, and reminded of the motorway driving route. For safety and legal reasons, two experimenters and one safety driver accompanied drivers throughout the study. The safety driver was seated in the passenger's seat and had access to an additional steering wheel and pedals, to intervene in the event of an emergency. The participants were asked to respect the rules of the highway code during manual driving, and keep a safe distance to surrounding traffic participants. Additionally, they were told that the vehicle was equipped with internal cameras for recording oral statements, feet movements, facial expressions, and the frontal and rear driving scene.

Prior to each experimental drive, participants were given the opportunity to practice driving (3-4 km on a rural road), allowing familiarisation with the vehicle and the ADF, after which they began the experimental drive.

The automated driving mode became available in the motorway if the following three criteria were all fulfilled: The AV:

- (1) was located in the centre of the lane,
 - (2) had a certain safety margin to the leading vehicle,
- and
- (3) was driving at less than 110 kph.

At that point, the vehicle dashboard turned blue, and the message 'the vehicle is ready for automated mode' was presented, along with an auditory alert. If the criteria for AV availability were not fulfilled, the experimenter instructed the participant to adjust the missing parameters. In order to hand over the driving task to the automated system, the participant was asked to release the acceleration pedal, and then push the 'R' button on the steering wheel. Once activated, the dashboard turned gold, and a sound was provided, which informed participants that they had activated the automated driving mode. During automated driving, participants were given different instructions regarding the range of activities that were allowed during automation, depending on the drive number.

In the first experimental drive, participants were instructed to hand over control as soon the automated driving mode was available, but they were always free to take over if they wanted to. While the automation was on, they were told that they could do whatever they liked, including engaging in an NDRT.

During the second experimental drive, the participants were asked to drive one half of the motorway section manually, and activate the automated mode during the other half of the drive. There were no instructions regarding a secondary task engagement for the period of automated driving in this drive.

Finally, in the third experimental drive, they were instructed to hand over control as soon as the automated driving mode was available, and they were encouraged to engage in an NDRT, such as reading a book, or playing on a smartphone. Immediately after the third and final drive, they completed the post-drive questionnaire, which incorporated questions on attitudes towards automation, and sensation seeking (not reported here). Finally, drivers were interviewed, and asked how the automated vehicle influenced their behaviour.

Across all drives, participants were prompted by the AV to take-over manual control of the vehicle one minute before the motorway exit, or 10 seconds prior to an unexpected event – in these situations the message '*You have 60 s (or 10 s) to take over control*' was displayed on the dashboard, accompanied by an auditory cue. To take over the driving task, the participants had to press the button 'O' on the steering wheel, or press the acceleration pedal, or turn the steering wheel.

3. Results and Discussion

Data analysis for this study is currently in progress, and results will be reported in time for the conference in October. To understand how different traffic conditions affected engagement in NDRTs, the number and duration of glances to the road centre area will be compared across three different driving speeds: 0-40 kph, 40-80 kph, and 80-120 kph, to understand how engagement in NDRTs is different for low speed of travel during heavy traffic, compared to automated driving at higher speeds. The correlation between mean speed and standard deviation of speed, and glance behaviours (number and duration of glances) will also be analysed. The effect of speed fluctuations on engagement in NDRT will be investigated, by setting acceleration thresholds for predicting glance presence from an accumulated frequency curve.

The findings of this study will be used to understand drivers' propensity to engage in NDRTs during real-world L3 automated driving, and how this is impacted by the vehicle's travelling speed, and presence of surrounding traffic. The implications of these results in terms of the design of more supportive HMI, and how different time budgets for transition of control affect behaviour in real world automated driving will be discussed.

References

- Christoph, M., Wesseling, S., & van Nes, N. (2019). Self-regulation of drivers' mobile phone use: The influence of driving context. *Transportation research part F: traffic psychology and behaviour*, 66, 262-272.
- Funkhouser, D., & Sayer, J. (2012). Naturalistic census of cell phone use. *Transportation research record*, 2321(1), 1-6.
- Goncalves, R. C., Louw, T. L., Quaresma, M., Madigan, R., & Merat, N. (2020). The effect of motor control requirements on drivers' eye-gaze pattern during automated driving. *Accident Analysis & Prevention*, 148, 105788.
- Louw, T., Kuo, J., Romano, R., Radhakrishnan, V., Lenné, M. G., & Merat, N. (2019). Engaging in NDRTs affects drivers' responses and glance patterns after silent automation failures. *Transportation research part F: traffic psychology and behaviour*, 62, 870-882.
- Morando, A., Gershon, P., Mehler, B., & Reimer, B. (2021). A model for naturalistic glance behavior around Tesla Autopilot disengagements. *Accident Analysis & Prevention*, 161, 106348.
- Noble, A. M., Miles, M., Perez, M. A., Guo, F., & Klauer, S. G. (2021). Evaluating driver eye glance behavior and secondary task engagement while using driving automation systems. *Accident Analysis & Prevention*, 151, 105959.
- Tivesten, E., & Dozza, M. (2015). Driving context influences drivers' decision to engage in visual-manual phone tasks: Evidence from a naturalistic driving study. *Journal of safety research*, 53, 87-96.
- Zeeb, K., Buchner, A., & Schrauf, M. (2016). Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident analysis & prevention*, 92, 230-239.

Brief Online Mindfulness Training in Young Drivers: A Pilot Randomised Controlled Trial

Derek A. Albert^{1*}, Marie Claude Ouimet³, Thomas G. Brown^{1,2,3}

¹ Department of Psychiatry, McGill University,
845, Sherbrooke St W, Montreal, Quebec, H3A 0G4, Canada

² Douglas Hospital Research Centre,
6875, LaSalle Blvd, Montreal, Quebec, H4H 1R3, Canada

³ Current Affiliation: Faculty of Medicine and Health Sciences, University of Sherbrooke,
150, Charles-Le Moyne PL, Suite 200, Longueuil, Quebec, J4K 0A8, Canada

*derek.albert@mail.mcgill.ca

Abstract: Driver distraction is a leading contributor to road traffic crashes in young drivers. Mind wandering (MW), a form of distraction involving off-task thoughts, is linked to crashes and, in some cases, unsafe driving (e.g., speeding). Brief online mindfulness training (MT) may reduce crash risk by enhancing reflexive awareness (i.e., meta-awareness) of MW and reducing its occurrence. This pre-post (T1, T2), randomised, placebo-controlled, double-blinded pilot trial examined these proposed mechanisms of MT and clarified its feasibility, in terms of acceptability and adherence, in young drivers. Twenty-six drivers aged 21-25 were allocated to either brief online MT (experimental) or progressive muscle relaxation (PMR, control), lasting 4–6 days. A custom website conducted randomisation, delivered interventions, administered questionnaires, and objectively tracked adherence. At T1 and T2, participants drove in a simulator and indicated MW whenever they noticed it, to assess meta-awareness, and when prompted by a thought-probe, to assess overall MW. Results showed that MT reduced MW while driving in simulation. Adherence and attrition did not differ significantly between interventions, but MT participants reported greater difficulty following instructions. Results support reduction in MW as a mechanism by which MT may reduce crash risk in young drivers. This preliminary evidence, alongside encouraging online adherence and acceptability data, may warrant definitive efficacy and effectiveness trials of online MT.

1. Introduction

Young drivers, aged 16-25, are overrepresented in road traffic crashes (World Health Organization, 2018). Driver distraction is a leading contributor to crashes in this population (Guo et al., 2017). Mind wandering (MW), a form of distraction involving off-task thoughts, is linked to crashes (Galéra et al., 2012) and, in some cases, unsafe driving (e.g., speeding) (Yanko & Spalek, 2014). Hence, there is a need to address this potential threat to young drivers.

Mindfulness training (MT) may protect young drivers from MW. MT involves, “paying attention in a particular way: on purpose, in the present moment, and non-judgementally” (Kabat-Zinn, 1994, p. 4). Brief MT, lasting four sessions, can reduce MW in attention tasks (Rahl et al., 2017). Thus, brief MT may reduce MW while driving. MT may also increase reflexive awareness, or meta-awareness of MW (Brandmeyer & Delorme, 2021). Evidence suggests that meta-awareness may reduce MW-related unsafe driving (Albert et al., 2018; Cowley, 2013). Therefore, these mechanisms of MT may reduce crash risk, but they have yet to be explored in the driving context.

Delivering MT online is now commonplace (Gál et al., 2021). Online MT is relatively inexpensive and logistically simple to deploy (Andersson & Titov, 2014; Boggs et al., 2014). It also increases accessibility by minimizing travel and scheduling constraints (González-García et al., 2021). Low acceptability, adherence, and retention, may plague online MT, however (Mrazek et al., 2019), but these essential feasibility metrics have yet to be assessed in young drivers.

This pilot trial examined two mechanisms by which MT may reduce crash risk. It was hypothesized that MT: H1) increases meta-awareness; and H2) reduces MW while driving. This pilot also clarified the feasibility of brief online in terms of acceptability and adherence. Results may support future definitive trials.

2. Method

2.1 Participants and recruitment

Twenty-six healthy drivers aged 21–25 were recruited via social media and classified ads. Inclusion and exclusion criteria are listed in Table A.1 (see Appendix A). Screening took place online. Participants were compensated \$60 CAD. The Douglas Mental Health University Institute Research Ethics Board (IUSMD-19-10) approved all procedures.

2.2 Study design

This pilot trial used a pre-post (T1, T2), randomised, placebo-controlled, double-blinded design. Participants were randomly assigned to one of two brief online interventions: Mindfulness Training (experimental); or Progressive Muscle Relaxation (PMR; control). A custom website conducted randomisation, delivered interventions, tracked adherence, and administered questionnaires.

2.3 Brief Online Interventions

Figure 1 shows a timeline of intervention and testing procedures. Participants were assigned one intervention session per day, over 4–6 days. Using the study website, participants completed one lab session at T1, 2–4 remote (at

home) sessions between T1 and T2, and one lab session at T2. Sessions involved 15 minutes of recorded audio instructions. The website administered post-session questionnaires.

2.3.1 Mindfulness Training (experimental):

Participants were instructed to: focus attention on breathing sensations; focus attention on other body sensations; notice and silently label thoughts as MW; and disengage from MW by re-focusing attention. Participants heard reminders and elaborations of these instructions throughout each session. Instructions were based on previous scripts (Rahl et al., 2017).

2.3.2 Progressive Muscle Relaxation (control):

Participants were instructed to: establish a slow, even breath; focus attention on particular muscle groups (e.g., arms and hands); notice and release tension. Participants were guided once through the whole body, then instructed to cycle through the various muscle groups on their own. Instructions were based on previous scripts (Feldman et al., 2010).

2.5.2 Meta-Awareness:

Participants were instructed, in both drives, to press the MW steering wheel button whenever they caught themselves MW. Self-caught MW reflects meta-aware MW. Self-caught MW rates, after controlling for probe-caught MW, operationalized meta-awareness (Zanesco et al., 2016).

2.6 Feasibility Measures

2.6.1 Acceptability of Interventions:

Per-session rates of positive and negative experiences, enumerated from participant descriptions in post-session questionnaires, measured acceptability for each intervention.

2.6.2 Adherence to Interventions:

Website playback logs objectively measured adherence. Completed sessions over remote-session days, calculated by sample and group, adjusted for variation in remote-session days (based on scheduling).

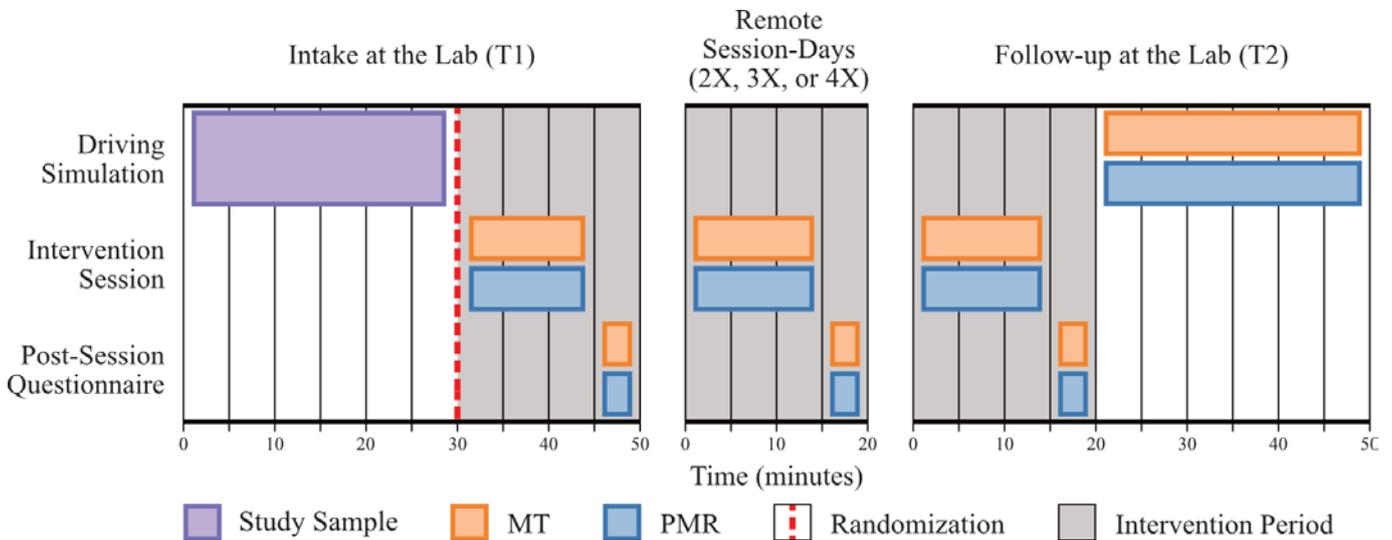


Figure 1. Timeline of procedures.

2.4 Driving Simulation

At T2 and T2, participants drove in a miniature University of Sherbrooke driving simulator (Brown et al., 2017). Each drive lasted 30 minutes. Participants drove on a circular (1 km radius), single carriageway road (90 km/h speed limit). Participants encountered oncoming traffic and a series of trucks traveling at 65 km/h in the ongoing lane, which partially obscured the oncoming lane. Participants were told to drive normally, which could include overtaking.

2.5 Outcome Measures

2.5.1 Mind Wandering:

Thought probes measured MW in both drives. Probe-tones, presented every 30–90 seconds, prompted participants to press one of two steering wheel buttons to indicate MW or focused driving. Probe-caught MW includes meta-unaware MW (Schooler et al., 2011). MW responses over total probes operationalized MW (Smallwood & Schooler, 2015).

3. Results

Participant demographic information can be found in Table A.2 (see Appendix A). Results for MW and meta-awareness are shown in Table 1. Comparing ΔT MW between groups revealed a significant difference. Only the MT group reported a significant decrease in MW. There was no significant between-group difference in ΔT meta-awareness.

Table 1 Results for Mind Wandering and Meta-awareness

| H | Variable | Contrast | | z | B | p |
|----|----------|----------|-------|-------|------|-----|
| | | Group | Time | | | |
| H1 | MW | MT-PMR | T2-T1 | -2.36 | 0.35 | .01 |
| | | MT | T2-T1 | -2.33 | 0.45 | .02 |
| | | PMR | T2-T1 | 0.86 | 1.27 | .39 |
| H2 | Meta | MT-PMR | T2-T1 | 1.19 | 1.42 | .12 |

Note. One-tailed contrasts yielded betas reflecting odds ratios from logistic mixed models of MW (H1), and rate ratios from Poisson mixed models of self-caught MW, or Meta (H2).

Regarding acceptability, rates of negative experiences differed significantly between groups, $z = 3.35$, $p = .001$, $RR = 5.16$. There were 25 negative experience reports over 31 descriptions among MT participants, and 5 over 32 among PMR participants. MT participants reported more difficulty following instructions. Positive experience rates did not differ significantly between groups.

Regarding adherence, participants completed 67.2% (43/64) of assigned remote intervention sessions. Adherence was 58% (18/31) in the MT group and 75.8% (25/33) in the PMR group. Overall, 19% (5/26) of participants were lost to follow-up (T2). Attrition was 30.8% (4/13) in the MT group and 7.69% (1/13) in the PMR group. Neither adherence nor attrition differed significantly between groups.

4. Discussion

As hypothesized, MT reduced MW while driving. In previous studies, online MT reduced MW in attention tasks (Bennike et al., 2017; Levinson et al., 2014). Our results suggest that online MT may also reduce MW while driving. This finding, in conjunction with recent evidence for fewer crashes in simulation following MT (Baltruschat et al., 2021), signals its promise for reducing MW and its potential consequences in real-world driving.

Results for meta-awareness were inconclusive. MT is proposed to cultivate meta-awareness that is sustained and non-propositional, reflecting continuous monitoring of thoughts (Dunne et al., 2019). Self-caught MW, which relies on intermittent, propositional judgements of one's mental state (e.g., "My mind is wandering!") (Schooler, 2002), may be insensitive to meta-awareness from MT. Using other measures of meta-awareness may clarify its role in MT effects on driver attention and behaviour.

MT was associated with more negative experiences in sessions than PMR. MT participants reported more difficulties, with statements such as, "I found it more difficult to...pay attention to the physical sensations in my body when there were longer periods of silence." MT practitioners may become frustrated or discouraged by frequent MW. This might explain why attrition is higher in MT compared to control conditions (Nam & Toneatto, 2016). Exploring methods to minimize frustration may boost retention in MT.

Intervention groups did not differ in attrition. Large group differences in attrition can indicate poor blinding (Hróbjartsson et al., 2014), variable intervention credibility (Alfonsson et al., 2016), and other confounds. Overall attrition in the present study was 19.2%, whereas average attrition across several in-person MT trials was found to be 29% (Nam & Toneatto, 2016). Future trials may identify and leverage features of online MT that contribute to higher retention of young drivers.

Intervention groups did not differ in adherence. Overall, 57.7% of participants completed all assigned sessions (ranging from 2–4). Forbes and colleagues (2018) reported 73.5% adherence to 4 sessions (out of 10 assigned, over 30 days, 10 minutes each). Low-intensity interventions (e.g., short, infrequent sessions) generally yield better adherence (Levensky et al., 2006), but high-intensity MT may be more effective (Strohmaier, 2020). Examining potential adherence costs and effectiveness benefits of different regimens may optimize MT for young drivers.

5. Conclusion

This study demonstrated a reduction in MW while driving from MT. This finding supports a mechanism by which MT may reduce young driver crash risk. Overall, this pilot trial reveals MT to be a feasible and compelling candidate for future definitive trials.

6. Acknowledgements

The authors would like to thank: Colin Courtney, Emily Freeman-Lavoie, and Michael Nolan for assisting with recording the audio MT and PMR sessions; Sam Watkinson for helping build the website that handled random assignment, adherence tracking, questionnaire administration, and intervention-delivery; Amedee d'Aboville for retrofitting the driving simulator with thought sampling capabilities; Lucie Legault and Lysiane Robidoux-Leonard for administrating and coordinating the study.

This study was supported by the Canadian Institutes of Health Research (MOP-137065, SAF-94813). The lead author was supported by scholarships from the Fonds de recherche du Québec – Santé (33885) and the Réseau de recherche en sécurité routière du Québec (677108). None of these funders were involved in: designing the study; collecting, analyzing or interpreting the data; writing the report; or the decision to submit the article for publication.

7. References

- Albert, D. A., Ouimet, M. C., Jarret, J., Cloutier, M.-S., Paquette, M., Badeau, N., & Brown, T. G. (2018). Linking mind wandering tendency to risky driving in young male drivers. *Accident Analysis & Prevention, 111*, 125–132. <https://doi.org/10.1016/j.aap.2017.11.019>
- Alfonsson, S., Olsson, E., & Hursti, T. (2016). Motivation and Treatment Credibility Predicts Dropout, Treatment Adherence, and Clinical Outcomes in an Internet-Based Cognitive Behavioral Relaxation Program: A Randomized Controlled Trial. *Journal of Medical Internet Research, 18*(3), e5352. <https://doi.org/10.2196/jmir.5352>
- Andersson, G., & Titov, N. (2014). Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry, 13*(1), 4–11. <https://doi.org/10.1002/wps.20083>
- Baltruschat, S., Mas-Cuesta, L., Cándido, A., Maldonado, A., Verdejo-Lucas, C., Catena-Verdejo, E., & Catena, A. (2021). Repeat Traffic Offenders Improve Their Performance in Risky Driving Situations and Have Fewer Accidents Following a Mindfulness-Based Intervention. *Frontiers in Psychology, 11*. <https://doi.org/10.3389/fpsyg.2020.567278>
- Bennike, I. H., Wieghorst, A., & Kirk, U. (2017). Online-based Mindfulness Training Reduces Behavioral Markers of Mind Wandering. *Journal of Cognitive Enhancement, 1*(2), 172–181. <https://doi.org/10.1007/s41465-017-0020-9>
- Boggs, J. M., Beck, A., Felder, J. N., Dimidjian, S., Metcalf, C. A., & Segal, Z. V. (2014). Web-Based Intervention in Mindfulness Meditation for Reducing Residual Depressive Symptoms and Relapse Prophylaxis: A Qualitative Study. *Journal of Medical*

- Internet Research*, 16(3), e87.
<https://doi.org/10.2196/jmir.3129>
- Brandmeyer, T., & Delorme, A. (2021). Meditation and the Wandering Mind: A Theoretical Framework of Underlying Neurocognitive Mechanisms. *Perspectives on Psychological Science*, 16(1), 39–66. <https://doi.org/10.1177/1745691620917340>
- Brown, T. G., Ouimet, M. C., Eldeb, M., Tremblay, J., Vingilis, E., Nadeau, L., Pruessner, J., & Bechara, A. (2017). The effect of age on the personality and cognitive characteristics of three distinct risky driving offender groups. *Personality and Individual Differences*, 113, 48–56. <https://doi.org/10.1016/j.paid.2017.03.007>
- Cowley, J. A. (2013). Off Task Thinking Types and Performance Decrements During Simulated Automobile Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1), 1214–1218. <https://doi.org/10.1177/1541931213571270>
- Dunne, J. D., Thompson, E., & Schooler, J. (2019). Mindful meta-awareness: Sustained and non-propositional. *Current Opinion in Psychology*, 28, 307–311. <https://doi.org/10.1016/j.copsyc.2019.07.003>
- Feldman, G., Greenson, J., & Senville, J. (2010). Differential effects of mindful breathing, progressive muscle relaxation, and loving-kindness meditation on decentering and negative reactions to repetitive thoughts. *Behaviour Research and Therapy*, 48(10), 1002–1011. <https://doi.org/10.1016/j.brat.2010.06.006>
- Forbes, L., Gutierrez, D., & Johnson, S. K. (2018). Investigating Adherence to an Online Introductory Mindfulness Program. *Mindfulness*, 9(1), 271–282. <https://doi.org/10.1007/s12671-017-0772-4>
- Gál, É., Ștefan, S., & Cristea, I. A. (2021). The efficacy of mindfulness meditation apps in enhancing users' well-being and mental health related outcomes: A meta-analysis of randomized controlled trials. *Journal of Affective Disorders*, 279, 131–142. <https://doi.org/10.1016/j.jad.2020.09.134>
- Galéra, C., Orriols, L., M'Bailara, K., Laborey, M., Contrand, B., Ribéreau-Gayon, R., Masson, F., Bakiri, S., Gabaude, C., Fort, A., Maury, B., Lemercier, C., Cours, M., Bouvard, M.-P., & Lagarde, E. (2012). Mind wandering and driving: Responsibility case-control study. *BMJ: British Medical Journal*, 345, e8105. <https://doi.org/10.1136/bmj.e8105>
- González-García, M., Álvarez, J. C., Pérez, E. Z., Fernández-Carriba, S., & López, J. G. (2021). Feasibility of a Brief Online Mindfulness and Compassion-Based Intervention to Promote Mental Health Among University Students During the COVID-19 Pandemic. *Mindfulness*, 12(7), 1685–1695. <https://doi.org/10.1007/s12671-021-01632-6>
- Guo, F., Klauer, S. G., Fang, Y., Hankey, J. M., Antin, J. F., Perez, M. A., Lee, S. E., & Dingus, T. A. (2017). The effects of age on crash risk associated with driver distraction. *International Journal of Epidemiology*, 46(1), 258–265. <https://doi.org/10.1093/ije/dyw234>
- Hróbjartsson, A., Emanuelsson, F., Skou Thomsen, A. S., Hilden, J., & Brorson, S. (2014). Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *International Journal of Epidemiology*, 43(4), 1272–1283. <https://doi.org/10.1093/ije/dyu115>
- Kabat-Zinn, J. (1994). *Wherever you go, there you are: Mindfulness meditation in everyday life*. Hyperion.
- Levensky, E. R., O'Donohue, W. T., & William, T. (2006). Patient adherence and nonadherence to treatments. *Promoting Treatment Adherence: A Practical Handbook for Health Care Providers*, 3–14.
- Levinson, D. B., Stoll, E. L., Kindy, S. D., Merry, H. L., & Davidson, R. J. (2014). A mind you can count on: Validating breath counting as a behavioral measure of mindfulness. *Frontiers in Psychology*, 5(OCT), 1–10. <https://doi.org/10.3389/fpsyg.2014.01202>
- Mrazek, A. J., Mrazek, M. D., Cherolini, C. M., Cloughesy, J. N., Cynman, D. J., Gougis, L. J., Landry, A. P., Reese, J. V., & Schooler, J. W. (2019). The future of mindfulness training is digital, and the future is now. *Current Opinion in Psychology*, 28, 81–86. <https://doi.org/10.1016/j.copsyc.2018.11.012>
- Nam, S., & Toneatto, T. (2016). The Influence of Attrition in Evaluating the Efficacy and Effectiveness of Mindfulness-Based Interventions. *International Journal of Mental Health and Addiction*, 14(6), 969–981. <https://doi.org/10.1007/s11469-016-9667-1>
- Rahl, H. A., Lindsay, E. K., Pacilio, L. E., Brown, K. W., & Creswell, J. D. (2017). Brief mindfulness meditation training reduces mind wandering: The critical role of acceptance. *Emotion (Washington, D.C.)*, 17(2), 224–230. <https://doi.org/10.1037/emo0000250>
- Schooler, J. W. (2002). Re-representing consciousness: Dissociations between experience and meta-consciousness. *Trends in Cognitive Sciences*, 6(8), 339–344. [https://doi.org/10.1016/S1364-6613\(02\)01949-6](https://doi.org/10.1016/S1364-6613(02)01949-6)
- Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, 15(7), 319–326. <https://doi.org/10.1016/j.tics.2011.05.006>
- Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology*, 66(1), 487–518. <https://doi.org/10.1146/annurev-psych-010814-015331>
- Strohmaier, S. (2020). The Relationship Between Doses of Mindfulness-Based Programs and Depression, Anxiety, Stress, and Mindfulness: A Dose-Response Meta-Regression of Randomized Controlled Trials. *Mindfulness*, 11(6), 1315–1335. <https://doi.org/10.1007/s12671-020-01319-4>
- World Health Organization. (2018). *Global status report on road safety 2018*. World Health Organization. <https://www.who.int/publications/i/item/9789241565684>
- Yanko, M. R., & Spalek, T. M. (2014). Driving with the wandering mind: The effect that mind-wandering has on driving performance. *Human Factors*, 56(2), 260–269. <https://doi.org/10.1177/0018720813495280>



Zanesco, A. P., King, B. G., MacLean, K. A., Jacobs, T. L., Aichele, S. R., Wallace, B. A., Smallwood, J., Schooler, J. W., & Saron, C. D. (2016). Meditation training influences mind wandering and mindless reading. *Psychology of Consciousness: Theory, Research, and Practice*, 3(1), 12–33.
<https://doi.org/10.1037/cns0000082>

Appendix A

Table A.1 Inclusion and Exclusion Criteria

| Criterion | Measure |
|--|---|
| Inclusion | |
| 1. Aged 21–25 | Self-report |
| 2. Valid driving license | Self-report, lab verification |
| 3. Normal or corrected vision and hearing | Self-report |
| 4. One or more years of independent driving | Self-report |
| Exclusion | |
| 1. Diagnosed neurological or psychiatric disorder | Self-report |
| 2. Generalized Anxiety Disorder symptoms | Total score > 10 on the Generalized Anxiety Disorders Questionnaire (Spitzer et al., 2006) |
| 3. Depression symptoms | Total score \geq 14 on the Beck Depression Inventory II (Beck et al., 1996) |
| 4. Alcohol Use Disorder symptoms | Total scores \geq 2 on items 4 and 6 of the Alcohol Use Disorders Questionnaire (Saunders et al., 1993; Johnson et al., 2013) |
| 5. Drug Use Disorder symptoms | Total scores \geq 2 on items 6 and 8 of the Drug Use Disorders Questionnaire (Berman et al., 2003; Hildebrand, 2015) |
| 6. Previous charge of driving while impaired | Self-report |
| 7. Meditation experience | Self-reported meditation practice \geq once per week in the past 6 months |
| 8. Personal or family history of Psychosis or Schizophrenia | Self-report |
| 9. Prodromal symptoms | Total score \geq 6 on the Prodromal Questionnaire (van der Gaag et al., 2012) |
| 10. Propensity to hyperventilate | Self-report |
| 11. Psychological trauma, recent bereavement, or personal crisis | Self-report |
| 12. Detectable blood alcohol | Alco-Sensor IV at lab |
| 13. Simulator sickness | Self-report following practice drive in simulator |

Note. Exclusion criteria: 1–7 controlled for factors affecting MW (Chen et al., 2019; Sayette et al., 2010; Smallwood, 2013); 8–11 minimized adverse effects from the interventions (Banks et al., 2015; Bernstein et al., 2007); 12 and 13 controlled for factors affecting driving behaviour.

Table A.2 Sample Demographics by Intervention

| Variable | MT (n = 13) | PMR (n = 13) |
|-------------------------------------|----------------|-----------------|
| Age, M (SD) | 23.8 (1.27) | 22.7 (1.01) |
| Sex, n male (%) | 7 (53.9) | 7 (53.9) |
| Ethnicity, n (%) | | |
| Other | 10 (76.9) | 7 (53.8) |
| Caucasian | 3 (23.1) | 6 (46.2) |
| Education level, n (%) | | |
| Some university | 8 (61.5) | 10 (76.9) |
| High school / college | 4 (30.8) | 3 (23.1) |
| Missing | 1 (7.70) | 0 (0.00) |
| Annual income, n (%) | | |
| \$6000 or more | 7 (53.8) | 8 (61.5) |
| \$0–\$5,999 | 6 (46.2) | 5 (38.5) |
| Employment, n (%) | | |
| Full-time studies + part-time work | 7 (53.8) | 5 (38.5) |
| Full-time work or full-time studies | 4 (30.8) | 6 (46.1) |
| Other | 2 (15.4) | 2 (15.4) |
| License type, n (%) | | |
| Probationary | 13 (100) | 11 (84.6) |
| Full | 0 (0.00) | 11 (84.6) |
| Traffic violations, n (%) | | |
| None | 11 (84.6) | 12 (92.3) |
| One in past 2 years | 2 (15.4) | 1 (7.69) |

Note. College refers to *Collège d'enseignement général et professionnel*, in Quebec, Canada. Full-time work \geq 35 hours/week. Probationary licenses can be obtained at \geq 17 years of age, following 12 months with a learner's license in Quebec. Drivers may obtain a full license after 2 years with a probationary license.

Body posture and physiological indicators for drowsiness detection in a partial automated driving.

Gaëtan Perrotte^{1,2*}, Arthur Portron¹, Clément Bougard^{1,2}, Jean-Louis Vercher¹

¹ Aix Marseille Univ, CNRS, ISM, Marseille, France

² Stellantis - Centre Technique de Vélizy, Vélizy-Villacoublay, France

*gaetan.perrotte@stellantis.com

Abstract: Driver monitoring is a crucial element for driving safety notably due to the large implementation of autonomous vehicles in the coming years. However, this monitoring must be rethought to adapt to new characteristics of such a mode. The purpose of this study is to examine the potential of associating postural and physiological information for monitoring drowsiness at the wheels. Twenty-two participants drove a static car simulator for 100 min in a monotonous environment, including 90 min in a level-2 autonomous mode. We measured physiological and behavioral indicators such as heart rate, respiration rate, eyelid movements (blink duration, frequency and PERCLOS) and body pressure distribution during the session. Psychomotor Vigilant Task-B (PVT-B) and Karolinska Sleepiness Scale (KSS) were also conducted before and after the driving session. Subjective perception of drowsiness was shown by a significant increase of KSS score between pre- and post-session measurements. Preliminary analyses of ocular data revealed large inter-individual variability in drowsiness level and timing during the session. Over half of the participants displayed high levels of drowsiness. Two participants were selected for further analysis. In these participants, elevated PERCLOS levels (> 50%) were correlated with decreased heart rate and center of pressure (COP) movements. Although further analysis needs to be performed, body pressure features could be a relevant information for the detection of drowsy driving.

1. Introduction

Drowsiness at the wheels, which can be considered as an inattention subcategory (Regan et al., 2011), is a major cause of death on roads especially on highways since it represents one third of fatal death accidents in France (ONISR, 2019). Sleep debt, sleep quality (Maia et al., 2013) or time of the day (Horne & Reyner, 1995) are key contributors to the apparition of drowsiness at the wheels. Driving environments and particularly monotonous roads are also factors of vigilance decrement and may lead to the apparition of microsleeps (Larue et al., 2011; Thiffault & Bergeron, 2003). Partial autonomous vehicles (i.e., level 2 & 3 defined by SAE; SAE International, 2016) have been developed to take part of the driving tasks in such situations with the development of “Highway driver assist” functionality. Nevertheless, following the definition provided by SAE, although the driver will not control the lateral and longitudinal movements of the car, he/she has to maintain a cognitive awareness on the driving environment and intervene in any safety critical event to which the automation would not act properly (McWilliams & Ward, 2021).

To make sure that the drivers could take-over their vehicle at any time, monitoring systems are used to detect altered states such as drowsiness. Nowadays, monitoring systems are mostly based on vehicular-behavioral data (e.g., lane deviation, lane crossing and steering-wheel angle) and drivers’ facial information (e.g., gaze, blink, percentage of eye closure (PERCLOS); see Halin et al., 2021 for a review). However, since in autonomy mode car movements are mostly controlled by the vehicle itself and drivers may

be engaged in non-driving related tasks (NDRTs), drivers will no longer be constantly facing the camera, making video monitoring irrelevant. New monitoring systems implementing alternative features must emerge. As an alternative to video systems, physiological data are increasingly being explored to detect drowsiness while driving, including the use of heart rate variability (Buendia et al., 2019; Fujiwara et al., 2019). However, Persson et al. (2020) demonstrated the difficulty to detect drowsiness based on this unique information during real car driving. Moreover, monitoring and recording heart activity in particular must interfere as least as possible with the driver’s activity and comfort, which is not compatible with classical methods using contact electrodes, and even wearable sensors. Other less invasive techniques should be preferred.

The use of seat-based sensors (e.g., Wusk & Gabler, 2018) could be a solution to provide such a usable (non-invasive) and relevant driver monitoring system. In addition to supplying physiological data, driver’s posture in the seat could be determined using pressure sensor matrices. However, little is known about seated posture and its potential link with drowsiness. In their study, Gwak et al. (2020) used data from the center of pressure (COP) distribution with a machine learning algorithm to detect drowsiness at the wheels and showed that this feature can be considered as promising. However, their algorithm also used eye data, which is known as the most relevant data for drowsiness detection (Schleicher et al., 2008).

The primary objective of the present study was to explore the potential of drowsiness detection and

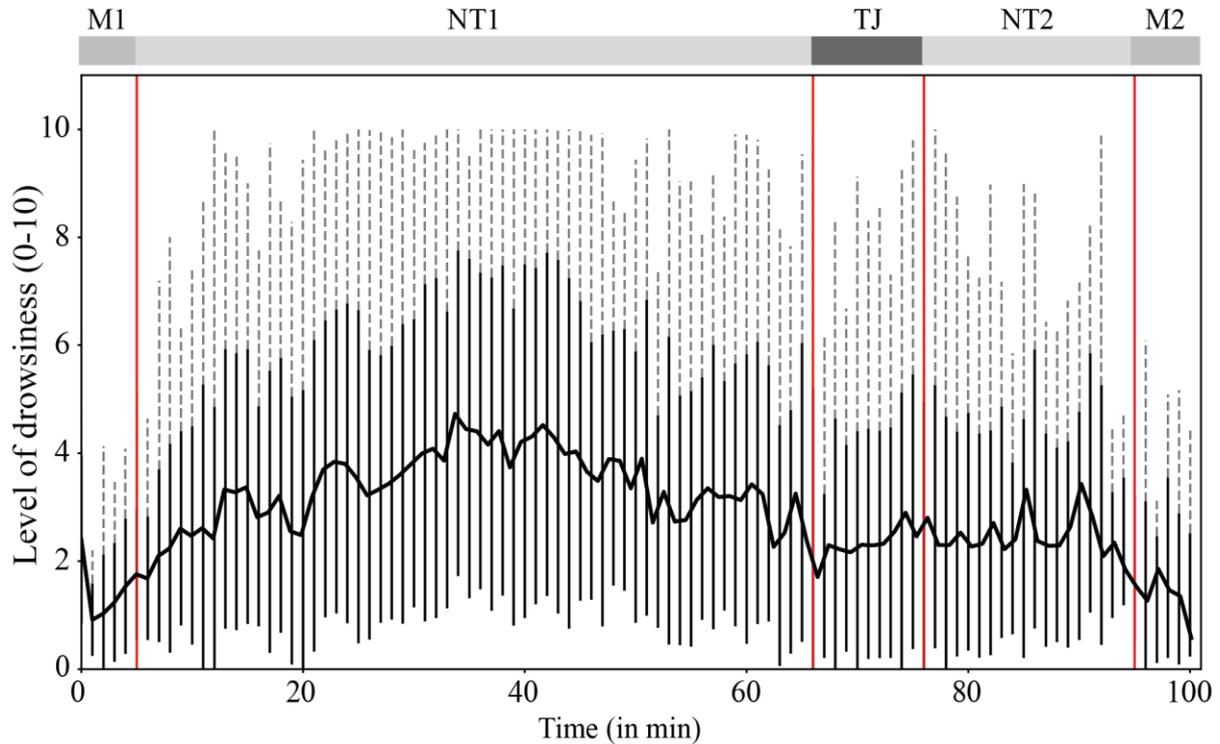


Figure 1 : Evolution of Drowsimeter R100 scores during the test session. Scores are averaged per minute for all participants. Mean value and standard deviation are represented by solid line. Dashed line represents the maximum value for each minute Autonomous driving regroupings NT1, TJ and NT2 periods. M = Manual; NT = No Traffic; TJ = Traffic Jam.

prediction by using body pressure features alone or coupled with physiological data. For that, we induced drowsiness during a prolonged automated driving session, with the aim of spotting features on physiological data and body pressure in relation with classical and already validated drowsiness indicators (eye aperture and movements).

2. Material and methods

2.1 Driving procedure

Twenty-two drivers (12 females, 10 males) aged between 19 and 31 years old took part in this study using a static driving simulator at the Mediterranean Center of Virtual Reality. They all received information about the experiment and agreed to participate. An ethical committee (agreement IRB00012476-2020-15-07-63) approved the protocol. On the day of the experiment, the participants were not allowed to smoke or to drink coffee or tea. After being informed about the study and the setup, they discovered the simulator and the autonomous mode during a 15-min familiarization phase. Then the test session was conducted in a Level-2 of automation: in this mode, the drivers were not allowed to engage themselves to NDRT which could maintain the drivers alert. This session proceeded as follows: participants drove under manual mode for 5 minutes (M1) at the end of which the system sent a manual-to-autonomous request to activate the “Highway driver assist”

function. On autonomous mode, all drivers had instructions to monitor the environment and to take-over the vehicle as soon as possible when requested by a visuo-auditory signal. The run on autonomous mode lasted 90 min by alternating three phases: (i) 60 min at 110 km/h without any traffic (NT1), (ii) 10 min of traffic-jam (TJ) and (iii) 20 min under the same conditions as the first one (NT2)(i). After that, a take-over request (TOR) was sent and an obstacle on the road had to be avoided. Once the vehicle had been taken back, drivers had to continue the travel for 5 min under manual mode (M2).

2.2 Data acquisition and processing

During the session, physiological data (i.e., heart rate, respiration rate and electrodermal activity) were recorded by using a BIOPAC® system MP150 at 500Hz. Body pressure distribution was recorded at 30 Hz with two textile pressure sensor mats (XSENSOR® Technology) placed on the seat (36 x 36 sensor cells) and on the backrest (64 x 40 sensor cells). The coordinates of the COP of each sensor mat was calculated on line. Ocular data (PERCLOS, blinks ...) were recorded by using Drowsimeter R100 (Phasya®), which uses a machine learning algorithm to obtain a drowsiness score [0 - 10] at 1 Hz (François et al., 2016). In addition, three video cameras were installed to record the participants (front, side, and feet).

Vigilance tests were also conducted before and after the driving test by using Karolinska Sleepiness

Scale (KSS, Akerstedt & Gillberg, 1990) and the Psychomotor Vigilance Task-B (PVT-B, Basner et al., 2011).

Heart rate was extracted from the raw ECG signal using Neurokit2 Python toolbox (Makowski et al., 2021). COP movement was obtained by calculating the Euclidian distance between the $(n-1)^{th}$ COP and n^{th} COP coordinates.

3. Preliminary results

3.1 Induction of drowsiness

Since monitoring a monotonous and poor environment during a long period of automation is known to induce vigilant decrement and drowsiness (Körber et al., 2015), we hypothesized that the longer the period in autonomous mode, the higher the level of drowsiness reached. Figure 1 shows the evolution of drowsiness scores during the test session. Scores are averaged per minute for all participants. As depicted, the first automated driving condition (~60 min) induced high levels of drowsiness, with in average a maximum at 33 min of automated driving, while the third sequence of automated driving (~20min) induced moderate levels of drowsiness. However, a significant inter-individual effect was observed on the dynamics of drowsiness onset during the session, as illustrated by the maximum score of drowsiness achieved over the minute (Fig 1, dashed line). The detailed analysis of ocular data for all participants highlights that half of them display high levels of drowsiness ($n = 11$).

Results of the KSS scores confirmed an effect in subjective drowsiness perception, as shown by a significant increase of KSS score between pre- (4.09 ± 1.27) and post-session measurements (6.55 ± 1.34) ($p < 0.05$).

3.2 Physiological and postural analysis

The temporal dynamics of physiological and behavioral data was analyzed to better understand the onset and occurrence of drowsiness episodes. Figure 2 shows the evolution of PERCLOS, COP movement on the seat and mean heart rate from two participants over the session. In this figure, time periods with a high percentage of PERCLOS ($> 50\%$; in red) correlate with a decrease in heart rate and COP movement. The analysis of front camera video confirms that these individuals were asleep at this time. Taken together, this information could be useful in detecting sleepy drivers in autonomous mode. However, further data analysis should be performed to examine the early stages of drowsiness.

4. Conclusion

Preliminary results confirm that long durations of autonomous driving can induce high levels of drowsiness on monotonous roads in drivers. Furthermore, these preliminary results tend to show the interest of using postural information coupled with physiological information to detect drowsiness while driving. However, further analysis is needed to determine whether this kind of result can be

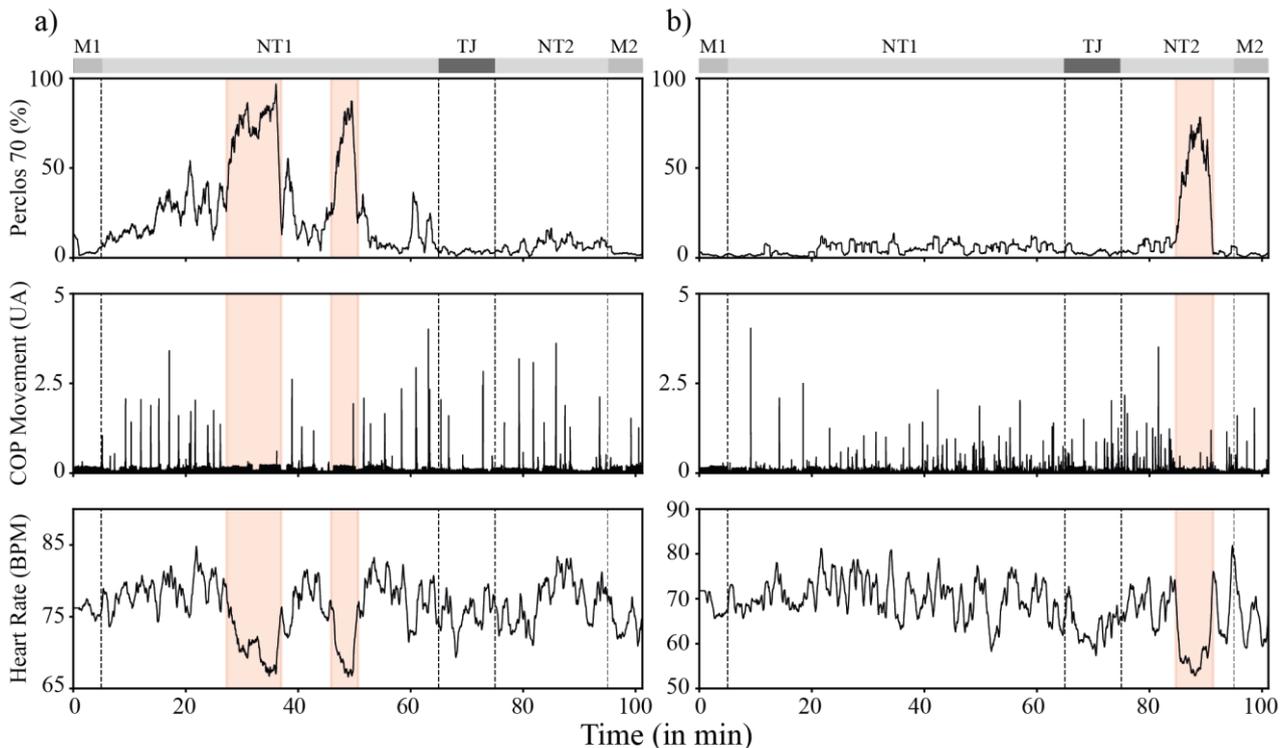


Figure 2 : Temporal dynamics of PERCLOS, COP movement on the seat cushion and heart rate during the test session for two participants. Zone in red represents periods with high-level of drowsiness (PERCLOS $> 50\%$). Autonomous driving regroups NT1, TJ and NT2 periods. M = Manual; NT = No Traffic; TJ = Traffic Jam.

generalized across all participants who felt asleep, but also to determine the correlation with earlier stages of drowsiness.

Acknowledgments

This study was part of the OpenLab agreement “Automotive Motion Lab” between Stellantis and Aix-Marseille University and CNRS.

5. References

- Akerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *The International Journal of Neuroscience*, 52(1-2), 29-37. <https://doi.org/10.3109/00207459008994241>
- Basner, M., Mollicone, D., & Dinges, D. F. (2011). Validity and Sensitivity of a Brief Psychomotor Vigilance Test (PVT-B) to Total and Partial Sleep Deprivation. *Acta astronautica*, 69(11-12), 949–959. <https://doi.org/10.1016/j.actaastro.2011.07.015>
- Buendia, R., Forcolin, F., Karlsson, J., Arne Sjöqvist, B., Anund, A., & Candefjord, S. (2019). Deriving heart rate variability indices from cardiac monitoring—An indicator of driver sleepiness. *Traffic Injury Prevention*, 20(3), 249-254. <https://doi.org/10.1080/15389588.2018.1548766>
- François, C., Hoyoux, T., Langohr, T., Wertz, J., & Verly, J. G. (2016). Tests of a New Drowsiness Characterization and Monitoring System Based on Ocular Parameters. *International journal of environmental research and public health*, 13(2), 174. <https://doi.org/10.3390/ijerph13020174>
- Fujiwara, K., Abe, E., Kamata, K., Nakayama, C., Suzuki, Y., Yamakawa, T., Hiraoka, T., Kano, M., Sumi, Y., Masuda, F., Matsuo, M., & Kadotani, H. (2019). Heart Rate Variability-Based Driver Drowsiness Detection and Its Validation With EEG. *IEEE Transactions on Biomedical Engineering*, 66(6), 1769-1778. <https://doi.org/10.1109/TBME.2018.2879346>
- Gwak, J., Hirao, A., & Shino, M. (2020). An Investigation of Early Detection of Driver Drowsiness Using Ensemble Machine Learning Based on Hybrid Sensing. *Applied Sciences*, 10(8), 2890. <https://doi.org/10.3390/app10082890>
- Halin, A., Verly, J. G., & Van Droogenbroeck, M. (2021). Survey and Synthesis of State of the Art in Driver Monitoring. *Sensors*, 21(16), 5558. <https://doi.org/10.3390/s21165558>
- Horne, J.A. and Reyner, L.A. (1995), Driver sleepiness. *Journal of Sleep Research*, 4: 23-29. <https://doi-org.lama.univ-amu.fr/10.1111/j.1365-2869.1995.tb00222.x>
- Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving. *Procedia Manufacturing*, 3, 2403-2409. <https://doi.org/10.1016/j.promfg.2015.07.499>
- Larue, G. S., Rakotonirainy, A., & Pettitt, A. N. (2011). Driving performance impairments due to hypovigilance on monotonous roads. *Accident Analysis & Prevention*, 43(6), 2037-2046. <https://doi.org/10.1016/j.aap.2011.05.023>
- Maia, Q., Grandner, M. A., Findley, J., & Gurubhagavatula, I. (2013). Short and long sleep duration and risk of drowsy driving and the role of subjective sleep insufficiency. *Accident Analysis & Prevention*, 59, 618-622. <https://doi.org/10.1016/j.aap.2013.07.028>
- Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Schölzel, C., & Chen, S. H. A. (2021). NeuroKit2 : A Python toolbox for neurophysiological signal processing. *Behavior Research Methods*, 53(4), 1689-1696. <https://doi.org/10.3758/s13428-020-01516-y>
- ONISR. (2019). La sécurité routière en France Bilan de l'accidentalité de l'année 2019. Retrieved from https://www.onisr.securite-routiere.gouv.fr/sites/default/files/2020-09/Bilan_2019_version_site_internet_24_sept.pdf
- Persson, A., Jonasson, H., Fredriksson, I., Wiklund, U., & Ahlstrom, C. (2020). Heart Rate Variability for Classification of Alert Versus Sleep Deprived Drivers in Real Road Driving Conditions. *IEEE Transactions on Intelligent Transportation Systems*, 1-10. <https://doi.org/10.1109/TITS.2020.2981941>
- Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis & Prevention*, 43(5), 1771-1781. <https://doi.org/10.1016/j.aap.2011.04.008>
- SAE International. (2016). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. Retrieved from https://www.sae.org/standards/content/j3016_202104
- Schleicher, R., Galley, N., Briest, S., & Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings : Looking tired? *Ergonomics*, 51(7), 982-1010. <https://doi.org/10.1080/00140130701817062>
- Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue : A simulator study. *Accident Analysis & Prevention*, 35(3), 381-391. [https://doi.org/10.1016/S0001-4575\(02\)00014-3](https://doi.org/10.1016/S0001-4575(02)00014-3)
- Wusk, G., & Gabler, H. (2018). Non-Invasive Detection of Respiration and Heart Rate with a Vehicle Seat Sensor. *Sensors*, 18(5), 1463. <https://doi.org/10.3390/s18051463>

Data Augmentation via Neural-Style-Transfer for Driver Distraction Recognition

Che-Tsung Lin^{1*}, Thomas Streubel^{2,3}, Marco Dozza², Giulio Bianchi Piccinini², Christopher Zach¹

¹ Department of Electrical Engineering, Chalmers University of Technology, SE-41296, Gothenburg, Sweden

² Department of Mechanics and Maritime Sciences, Chalmers University of Technology, SE-41296, Gothenburg, Sweden

³ Current affiliation: Volvo Car Corporation AB, Safety Center, SE-40531, Gothenburg, Sweden

*chetsung@chalmers.se; alexofntu@gmail.com

Abstract: According to the National Highway Traffic Safety Administration, 3142 people were killed in motor vehicle crashes involving distracted drivers in 2019. Naturalistic driving datasets (NDD) have been widely used to study distracting activities while driving, with the aim of improving road safety. However, the time required to annotate videos to identify distracting activities is a major issue for research using NDD. Although full automation of the annotation process is not possible, the use of image classifiers is a way forward to hasten the classification of distractions and therefore the analysis of NDD. This paper presents the results obtained by applying image classifier to the publicly available Distracted Driver Dataset (DDD) and a sample of frames extracted from the EuroFOT and DriveC2X dataset. The results show that using ResNet-50 pretrained on Imagenet and Stylized Imagenet produces the highest accuracy on both DDD and our EuroFOT and DriveC2X datasets. The accuracy of the image classifier will now be tested on a different sample of the Swedish EuroFOT dataset, before using the image classifier for detecting distracting activities in other NDD. The faster identification of distracting activities will considerably hasten the future analyses of NDD.

1. Introduction

According to the National Highway Traffic Safety Administration (NHTSA) (National Highway Traffic Safety Administration, 2022), 3142 people were killed in motor vehicle crashes involving distracted drivers. Research on distraction has largely benefit from the analysis of Naturalistic Driving Data (NDD) in the last 15 years (for some examples, see Hickman & Hanowski, 2012; Klauer et al., 2006; Victor et al., 2015). However, one of the main concerns associated to the use of NDD is the time intensive and costly process of video reduction to extract variables from the videos. The complete automation of this process is hindered by different factors, such as the poor quality of videos and the lighting conditions. However, efforts have been conducted to perform video reduction of NDD through computer vision algorithms (see, for example Kuo et al., 2014).

Recently, Eraqi et al. (2019) made public the Distracted Driver Dataset (DDD) which provides images

exemplifying driver distraction behaviours especially related to phone usages such as phone talking/listening right, phone talking/listening left, texting right, and texting left (see Fig. 1). Eraqi et al. (2019) also applied Convolutional Neural Networks (CNNs) to perform image classification on DDD, with good preliminary results (81.69% using ResNet-50).

Convolutional Neural Networks (CNNs) are commonly thought to recognise objects by learning increasingly complex representations of object shapes. However, some recent studies suggested a more important role of image textures instead. Geirhos et al. (2019) has shown that CNNs pretrained by ImageNet (Deng et al., 2009) are strongly biased towards recognising textures rather than shapes, which is in stark contrast to human behavioural evidence and reveals fundamentally different classification strategies. Therefore, they further showed that a CNN can benefit from learning shape-based representation when trained on ‘Stylized ImageNet’, a stylized version of ImageNet. This is created by performing AdaIN style transfer (Huang et al., 2017) on the whole ImageNet dataset. Fig. 2



Fig. 1. Different driver's distracting activities from Distracted Driver Dataset

shows that the texture of the objects is no longer highly predictive of the target class, while the global shape tends to be retained. In this way, the texture and the shape of the objects are disentangled such that a CNN can be induced to learn from the shape instead of the texture.



Fig. 2. Visualisation of Stylized-ImageNet. Leftmost image: randomly selected ImageNet image of ring-tailed lemur. Others: neural-style-transferred counterparts where the texture cues are no longer highly predictive.

Overfitting and over-confidence are two major issues that easily arise when training CNNs. There are several regularization techniques in deep learning to address the former. For example, weight decay, early stopping, and dropout are some of the most popular ones. For the latter, model calibration such as temperature scaling (Guo et al., 2017), a single-parameter variant of Platt Scaling (Platt 1999), is proven to be effective. Label smoothing (Rafael et al., 2019) is a regularization technique that perturbs the target variable to make the model less certain of its predictions. It is viewed as a regularization technique because it restrains the largest logits fed into the softmax function from becoming much bigger than the rest. Moreover, the resulting model is found to be better calibrated. Therefore, the reason that label smoothing stands out is that it can deal with both issues at the same time. In recent years, image classification has been significantly improved by CNNs, such as Alexnet (Krizhevsky et al., 2012), GoogLeNet (Szegedy et al., 2015), VGG16 (Simonyan et al., 2015) and the powerful ResNets (He et al., 2016). These models kept pushing the limit of ImageNet Classification by providing better CNN structure. Then, the weightings of these models can be the initial values on a new problem in terms of transfer learning. i.e., a better pretrained CNN is helpful for the subsequent finetuning task such as driver distraction classification. In other words, the better the backbone, the higher the classification performance on DDD, EuroFOT, and Drive C2X. Label Smoothing (Rafael et al., 2019) is a regularization technique that introduces noise for the labels. This accounts for the fact that datasets may have some mistakes in them, so maximizing the likelihood of $\log p(y|x)$ might result in over-fitting. Assume for a small constant ϵ , the target value of the training label y is $1-\epsilon$ and $\epsilon/k-1$ for the target class and others, respectively. i.e., the original target value of each class is

$$P_i = \begin{cases} 1, & i = y, \\ 0, & i \neq y. \end{cases} \quad (1)$$

After label-smoothing, they become

$$P_i = \begin{cases} 1 - \epsilon, & i = y, \\ \epsilon/(k - 1), & i \neq y. \end{cases} \quad (2)$$

Therefore, for cross-entropy loss

$$Loss = -\sum_{i=1}^k p_i \log q_i, \quad (3)$$

the loss corresponding to each class is

$$Loss_i = \begin{cases} (1 - \epsilon) * Loss, & i = y, \\ \epsilon * Loss, & i \neq y. \end{cases} \quad (4)$$

This paper aims to present the results of applying state-of-the-art image classification techniques to two naturalistic driving datasets, to identify and categorize driver distraction tasks.

2. Method

As mentioned, image classification techniques were applied to two datasets: the DDD and a data sample extracted from the two European NDD collections EuroFOT (<https://www.eurofot-ip.eu/>) and Drive C2X (<https://cordis.europa.eu/project/id/270410>).

The authors downloaded the DDD from the website https://heshameraqi.github.io/distraction_detection. The dataset is split into training and validation sets and includes labelling of the distraction activities in the different frames (see Fig. 1). The labelling was used in this work as a ground truth for the classification algorithm. Both datasets (DDD and EuroFOT+Drive C2X) were split in training and validation sets, as indicated in Table 1 and Table 2, respectively. Initially, different image classification techniques were applied to the DDD.

Table 1 Number of training and validation frames for each distraction activity in Distracted Driver Dataset

| Class and images | Training | Validation |
|----------------------|----------|------------|
| Safe Driving | 2640 | 346 |
| Phone Right | 1505 | 213 |
| Phone Left | 1062 | 194 |
| Text Right | 945 | 180 |
| Text Left | 1150 | 170 |
| Adjusting Radio | 953 | 170 |
| Drinking | 933 | 143 |
| Hair or Makeup | 891 | 143 |
| Reaching Behind | 898 | 146 |
| Talking to Passenger | 1579 | 218 |

Table 2 Number of training and validation frames for each distraction activity in EuroFOT and Drive C2X sample

| Class and images | Training | Validation |
|-------------------------------|----------|------------|
| No activities | 169213 | 200432 |
| Interaction with passenger | 1244 | 504 |
| Talking or singing | 18106 | 732 |
| Reaching for an object | 19232 | 8046 |
| Interaction with center stack | 6836 | 3439 |
| Eating/Drinking | 4975 | 2784 |
| Hands-face interaction | 29847 | 19598 |
| Reading | 1920 | 1058 |

In this work, we use ResNet-50 for performing distraction recognition on both datasets.

3. Results

Our experimental results were mainly done by using ResNet-50 pretrained on ImageNet and Stylized ImageNet. Fig. 3 shows a sample image of the stylized DDD which presents the same image transformed by different style images. When we mix original DDD and stylized DDD, as can be seen in Table. 3, the accuracy is further boosted.

Besides, label smoothing is also quantitative beneficial in all cases.

The image classification technique which had very good performance for the DDD (88.05% see Table 3) was also applied to the labelled dataset obtained from the projects EuroFOT and Drive C2X. The resulting accuracy was 84.72%.

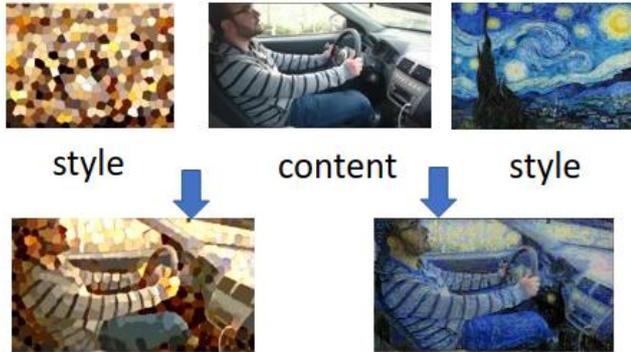


Fig. 3. Visualisation of Stylized-DDD

Table 3 ResNet-50 results on Distracted Driver Dataset

| Pretraining | Finetuning | Label smoothing | Accuracy |
|------------------------------|-------------------|-----------------|------------------------------|
| ImageNet | DDD | No | 81.695% (Eraqi et al., 2019) |
| ImageNet | DDD | Yes | 86.95% |
| ImageNet + Stylized-ImageNet | DDD | Yes | 88.05% |
| ImageNet + Stylized-ImageNet | DDD+ Stylized-DDD | Yes | 89.01% |
| ImageNet + Stylized-ImageNet | EuroFOT | Yes | 84.72% |

4. Discussion and conclusions

In this work, we explored the idea of applying neural-style transfer, to identify different distracting activities from images. The results show that ResNet-50 trained by ImageNet and Stylized-ImageNet can boost the performance on the target task. For DDD, we found that mixing DDD and the Stylized-DDD together with label smoothing can achieve the highest performance (7.306% higher accuracy over the results reported by Eraqi et al. (2019), using the same CNN). For our EuroFOT and Drive C2X datasets, we also found that the same model leads us to very competitive results. Both results clearly show that our image classifier works well on staged and real-driving datasets.

The next step is to validate the image classifier using a larger sample of the EuroFOT dataset. This sample will contain frames extracted from videos recording drivers who were not included neither in the training nor in the validation sets presented in Table 2. Based on the results of the validation, further developments of the image classifier might be required. Once the image classifier has reached the highest possible accuracy, analyses of naturalistic driving dataset involving distracting activities (see Ismaeel et al. [2020].

Morgenstern et al. [2020] and Tivesten et al. [2014] for some examples of this type of analyses) will be significantly hastened, due to less time dedicated to manual annotations. Furthermore, the use of this image classifier will enable analyses that were previously discarded due to the overbearing effort required for data reduction by manual annotations.

5. Acknowledgments

The authors are grateful to Vinnova and to the Chalmers Transport Area of Advance that sponsored this research. Vinnova funded this work through the project “Field Operational Test dataset enrichment for the development of safety and highly autonomous vehicle functions” (FOT-e). Chalmers Transport Area of Advance supported this research through the project “AoA”.

References

- Deng, J., et al. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255).
- Eraqi, H. M., Abouelnaga, Y., Saad, M. H., & Moustafa, M. N. (2019). Driver distraction identification with an ensemble of convolutional neural networks. *Journal of Advanced Transportation*, 2019.
- Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F. A., & Brendel, W. (2018). ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. *arXiv preprint arXiv:1811.12231*.
- Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. In *International Conference on Machine Learning* (pp. 1321-1330).
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Hickman, J. S., & Hanowski, R. J. (2012). An assessment of commercial motor vehicle driver distraction using naturalistic driving data. *Traffic injury prevention*, 13(6), 612-619.
- Huang, X., & Belongie, S. (2017). Arbitrary style transfer in real-time with adaptive instance normalization. In *Proceedings of the IEEE international conference on computer vision* (pp. 1501-1510).
- Ismaeel, R., Hibberd, D., Carsten, O., & Jamson, S. (2020). Do drivers self-regulate their engagement in secondary tasks at intersections? An examination based on naturalistic driving data. *Accident Analysis & Prevention*, 137, 105464.
- Klauer, C., 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*. Report DOT HS 810 594.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- Kuo, J., Koppel, S., Charlton, J. L., & Rudin-Brown, C. M. (2014). Computer vision and driver distraction: Developing a behaviour-flagging protocol for naturalistic

- driving data. *Accident Analysis & Prevention*, 72, 177-183.
- Morgenstern, T., Petzoldt, T., Krems, J. F., Naujoks, F., & Keinath, A. (2020). Using European naturalistic driving data to assess secondary task engagement when stopped at a red light. *Journal of safety research*, 73, 235-243.
- National Highway Traffic Safety Administration (2022). *Distracted Driving*. Retrieved March 29, 2022, from <https://www.nhtsa.gov/risky-driving/distracted-driving>.
- Platt, J. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, 10(3), 61-74.
- Szegedy, C., et al. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- Tivesten, E., & Dozza, M. (2014). Driving context and visual-manual phone tasks influence glance behavior in naturalistic driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 26, 258-272.
- Victor, T., et al. (2015). *Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk*. SHRP 2 Report S2-S08A-RW-1.

AutoConduct: a novel dataset for in-vehicle driver posture monitoring

Mingming Zhao¹, Georges Beurier¹, Hongyan Wang², Xuguang Wang^{1*}

¹ Univ Lyon, Univ Gustave Eiffel, Univ Claude Bernard Lyon 1, LBMC UMR_T9406, F69622, Lyon, France

² School of Automotive Studies, Tongji Univeristy, 4800 Cao An road, 201804, Shanghai, China

*xuguang.wang@univ-eiffel.fr

Abstract: Posture monitoring of vehicle drivers has a wide variety of applications, e.g., the detection of driver distraction and the identification of driver seated position, which can be integrated into automotive systems to improve active and passive safety. In spite of the advancement in sensing technologies and artificial intelligence, the research on the posture estimation better suited to in-car applications is far behind, partly because of lack of posture datasets with complete and accurate data annotations. To facilitate the progress into more realistic driver posture monitoring, this paper presents a large in-vehicle driver posture dataset called AutoConduct. The raw data, measured by 3 depth cameras and 2 pressure pads, were collected on a laboratory mockup from 23 drivers performing 42 activities including driving and non-driving related tasks. The ground truth driver postures were reconstructed from markers attached on the body measured by a motion capture system. The measurements from different sensors were temporally synchronized and spatially aligned. The data augmentation pipeline, based on computer graphics techniques, allows automatic generation of a great number of synthesized images with 2D and 3D annotations. With help of this dataset, several learning algorithms were tested to estimate the 3D skeleton including the position of head, trunk, arms, shanks and feet from the measurements of depth cameras or pressure sensors. The proposed monitoring functions demonstrated that the proposed framework could be used to develop more performant driver posture monitoring systems useful for the detection of driver distraction.

1. Introduction

Driver distraction has been reportedly regarded as one of the influential contributing factors of road traffic accidents (Beanland et al., 2013; Née et al., 2019). With the development of driving automation, counter-measures of driver distraction remain as important as ever, because current automation technologies potentiate out-of-the-loop problems if the driver is engaged in non-driving related tasks (Lee et al., 2021; Lu et al., 2016; Yoon & Ji, 2019). A Driver Posture Monitoring System (DPMS) can provide fundamental information for evaluating driver’s attention (Deo & Trivedi, 2019; Hu et al., 2020; Venturelli et al., 2017; Xing et al., 2017). In addition, the tracking of driver body locations can be used to modulate the collision response of restraint systems for better protection (Filatov et al., 2019).

Although a vast body of vision-based DPMS have been proposed (Wang et al., 2019), the performance of these systems remain limited because of the challenges present in this context such as the close proximity between body and vehicle interior, body occlusions and suboptimal camera placement. The research is further delayed by lack of in-vehicle posture datasets required by the supervised posture learning models.

Recently, the creation of in-vehicle driver posture datasets has become a trending topic (Borges et al., 2021; Borghi et al., 2017; Feld et al., 2020; Roth & Gavrilu, 2019), because the datasets allow the direct comparison of different methods with state of the art and stimulating research community. Nevertheless, the existing datasets are subject to the incompleteness of data annotations and body part coverage. In addition, few datasets are publicly available.

In order to develop robust driver posture monitoring systems, we present a novel framework to create a well-structured and extensive in-vehicle dataset, named AutoConduct. Based on this dataset, several monitoring functions are proposed and tested.

2. AutoConduct dataset

2.1 Data collection

Twenty-three drivers (11 females) with different age, height and BMI were asked to perform 42 driving and non-driving related tasks on a mockup. These tasks were extracted from previous studies (Dingus et al., 2006; Naujoks et al., 2018) to cover a range of in-vehicle posture variations. A detailed list of these tasks can be found in Zhao et al., (2021a)

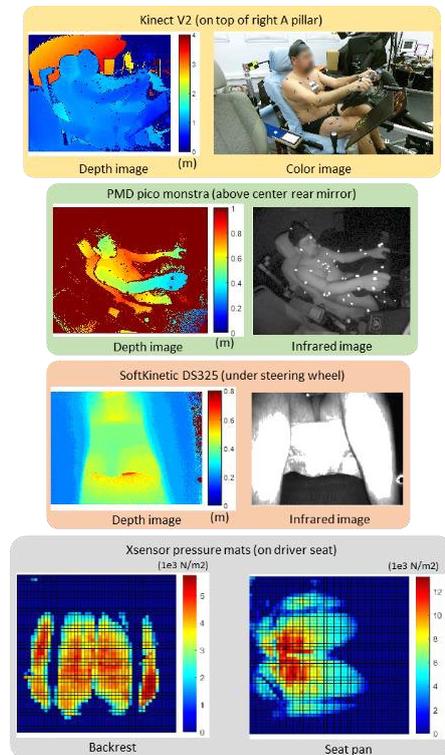


Fig. 1. Real data collected from experiment

The postural measurement (Fig. 1) was composed of multiple image flows from three depth cameras placed at different positions and body pressure distributions from two Xsensor pressure mats on seat pan and backrest. Meanwhile, the driver motion was recorded by an optical motion capture system VICON. Measurements from different sensors were electronically synchronized.

2.2 Data processing

Driver motions recorded by VICON system were reconstructed using RPx (Monnier & Wang, 2009), where the body posture was represented by joint angles or 3D joint positions and served as ground truth. Standard calibration method (Zhang, 2002) was performed to obtain the intrinsic and extrinsic parameters of each depth camera. Then we converted the depth images to point clouds which were then aligned with the reconstructed ground truth posture in the same world coordinate system, as shown in Fig. 2.

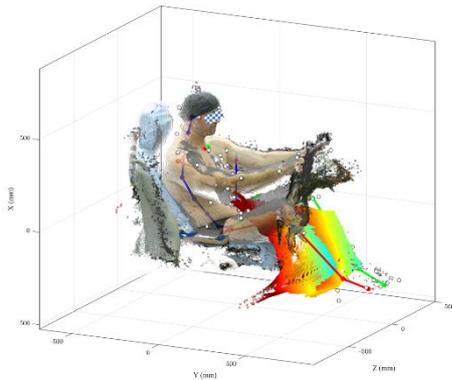


Fig. 2. Spatial alignment between point clouds from depth cameras and ground truth posture from VICON

2.3 Data augmentation

Inspired by previous studies (Cruz et al., 2020; Martinez-Gonzalez et al., 2020; Shotton, 2011), we established a pipeline using computer graphic techniques to enrich driver posture data samples for machine learning or deep learning algorithms (Zhao et al., 2020). The basic idea was to animate rigged virtual human characters with realistic external envelopes including skin, clothes, hair and predefined body part labels using reconstructed motions (Fig. 3). This allowed us to synthesize artificial images by rendering and meanwhile provides ground truth labels including body part segmentations and 3D skeleton.

2.4 AutoConduct vs state-of-the-art

Using the proposed framework, the AutoConduct dataset created in the present work consisted of two sets: real

data and synthetic data. The real data included ~130K frames of postural measurement including images and pressure distributions along with true driver postures. The synthetic dataset consisted of ~12 million data frames generated in a simulated in-vehicle scenario. Each frame included images and annotations.

In terms of the real dataset, the main advantage over the others (Borges et al., 2021; Borghi et al., 2017; Feld et al., 2020) is the availability of accurate ground truth postures in addition to the postural measurement covering the whole body, which enables the accuracy evaluation of posture estimation models and allows for investigations of posture monitoring approaches. Furthermore, the inclusion of body pressure distributions may be of help to provide complementary cues for robust posture monitoring.

Regarding the synthetic dataset, the proposed data generation pipeline allows one to automatically introduce variability in human shapes, body pose, background and view point configurations. Compared to the traditional data annotation and augmentation strategies (Torres et al., 2019; Yuen & Trivedi, 2018), the annotation labels can be automatically obtained almost for free, allowing one to scale up supervised learning to large scales.

3. Monitoring functions

3.1 Vision-based monitoring

3D body pose estimation: As opposed to 2D body pose, 3D pose facilitates understanding of driver activity, while the estimation of driver's 3D body pose is rarely investigated. Inspired by previous studies (Cao et al., 2019; Shotton et al., 2013), we proposed a posture estimation model based on body part localization and offset joint regression to extract 3D upper-body joints using a depth camera Kinect v2. The model was retrained on the synthetic dataset and evaluated on the real dataset. To reduce estimation errors caused by model uncertainty and body occlusions, a data-driven method (Plantard et al., 2017) was adapted to obtain more natural and more accurate driver postures. The percentage of the data frames with a predicted joint within 5 mm from the ground truth was 91% on average across the upper-body joints. This monitoring function will allow one to accurately identify whether driver's upper body is in normal position and whether driver's hands are on the steering wheel.

Head pose estimation: To predict the orientation and position of driver's head, a feature-based method was employed. Specifically, 3D facial keypoints were first extracted from the images of a depth camera using OpenPose (Cao et al., 2019). Then a Random Forest regression method learned from the true driver head motions was proposed to infer the head pose based on the keypoints positions. The balanced mean errors were less than 11° and 2 cm

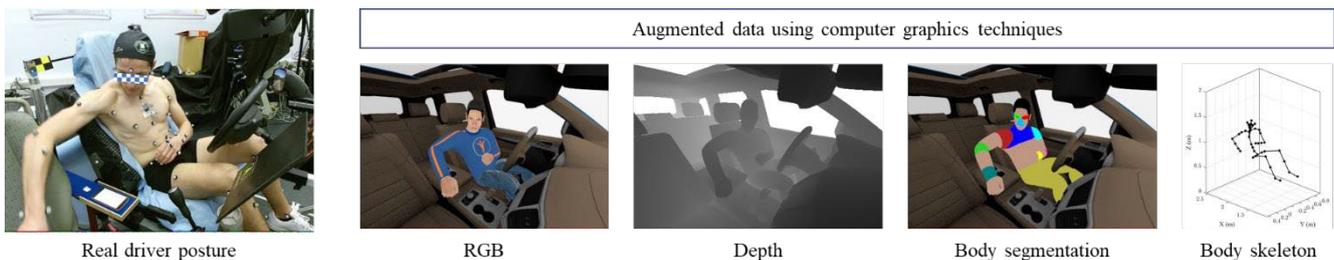


Fig. 3. Augmentation of in-vehicle driver posture images along with body part segmentation labels and 3D skeleton

respectively for the head orientation and position in 96.3% of the cases.

Shank pose estimation: The monitoring of driver's lower limbs is useful for the detection of pedal errors such as pedal misapplication. To this end, a shank posture estimation model based on clustering analysis was proposed to extract key points from the point cloud in the leg room. Machine learning classifiers were trained on the shank keypoints to predict left positions and right foot positions with an average accuracy of 93% and 88%, respectively.

The proposed monitoring functions were tested on a similar but different experiment setting, as shown in Fig. 4.

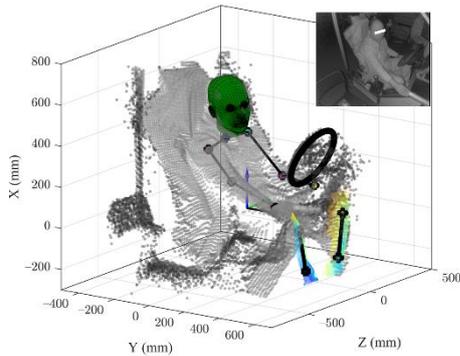


Fig. 4. Estimation of driver's posture using proposed vision-based monitoring functions

3.2 Pressure sensors based monitoring

In addition to the vision-based methods, driver posture classification in terms of different trunk and feet positions based on pressure distribution features were systematically investigated (Zhao et al., 2021b). Results showed that pressure sensors embedded into driver seat could provide reliable postural classification especially for normal and abnormal trunk positions, implying that pressure sensors could serve as good supplementary to cameras.

4. Conclusions

In this paper, we presented a procedure to create an in-vehicle driver posture dataset including data collection, processing and augmentation. The dataset showed advantages over existing ones regarding the coverage of body parts, data modalities, posture variations, the quality and completeness of data annotations. Based on this dataset, various monitoring functions were proposed to estimate driver's full-body posture in 3D, which allows one to identify if driver's hands, head, trunk and feet are in a non-driving position which may imply driver distraction. To the best of our knowledge, this is the first work that attempts to predict driver's full-body posture in 3D while providing quantified errors.

In the future, we will keep improving this procedure particularly the data augmentation pipeline. The synthetic dataset will be made open-access and more posture estimation models will be benchmarked on our dataset. Further effort is also needed to investigate the alignment between real in-vehicle data and the synthesized data. Another research direction is to explore sensor fusion methods to take advantage of multiple measurement inputs for better monitoring performance.

The driver's posture monitored by the proposed system could be a necessary input to improve passive safety

systems. It could also provide postural information to evaluate driver's distraction level. To this end, non-driving postures crucial to driving safety need to be classified and critical postural indicators useful for evaluating driver's attention need to be identified.

5. Acknowledgments

This research was partly funded by the French National Project ANR AutoConduct, grant number ANR-16-CE22-0007. The first author (M. Zhao) had a PhD scholarship from both the China Scholarship Council (n°201806260131) and Univ-Eiffel. The authors thank Ilias Theodorakos for participation in data collection, as well as Richard Roussillon, Stéphane Ardizzone, and Fabien Moreau for their technical assistance.

References

- Beanland, V., Fitzharris, M., Young, K. L., & Lenné, M. G. (2013). Driver inattention and driver distraction in serious casualty crashes: Data from the Australian National Crash In-depth Study. *Accident Analysis and Prevention*, 54, 99–107.
<https://doi.org/10.1016/j.aap.2012.12.043>
- Borges, J., Queirós, S., Oliveira, B., Torres, H., Rodrigues, N., Coelho, V., Pallauf, J., Brito, J. H. H., Mendes, J., & Fonseca, J. C. (2021). A system for the generation of in-car human body pose datasets. *Machine Vision and Applications*, 32(1).
<https://doi.org/10.1007/s00138-020-01131-z>
- Borghini, G., Venturelli, M., Vezzani, R., & Cucchiara, R. (2017). POSEidon: Face-from-Depth for driver pose estimation. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 5494–5503.
<https://doi.org/10.1109/CVPR.2017.583>
- Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2019). Realtime multi-person 2D pose estimation using part affinity fields. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua(Xxx)*, 1302–1310.
<https://doi.org/10.1109/CVPR.2017.143>
- Cruz, S. D. da, Wasenmuller, O., Beise, H. P., Stifter, T., & Stricker, D. (2020). SVIRO: Synthetic vehicle interior rear seat occupancy dataset and benchmark. *Proceedings - 2020 IEEE Winter Conference on Applications of Computer Vision, WACV 2020*, 962–971.
<https://doi.org/10.1109/WACV45572.2020.9093315>
- Deo, N., & Trivedi, M. M. (2019). Looking at the Driver/Rider in Autonomous Vehicles to Predict Take-Over Readiness. *IEEE Transactions on Intelligent Vehicles*, 5(1), 41–52.
<https://doi.org/10.1109/TIV.2019.2955364>
- Dingus, T. a., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. a., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., & Knippling, R. R. (2006). The 100-Car naturalistic driving study phase II – Results of the 100-Car field experiment. *Dot Hs 810 593, April*, No. HS-810 593.
<https://doi.org/DOT HS 810 593>
- Feld, H., Mirbach, B., Katrolia, J., Selim, M., Wasenmüller, O., & Stricker, D. (2020). DFKI Cabin Simulator: A

- Test Platform for Visual In-Cabin Monitoring Functions. *ArXiv:2002.03749v2*.
<http://arxiv.org/abs/2002.03749>
- Filatov, A., Scanlon, J. M., Bruno, A., Danthurthi, S. S. K., & Fisher, J. (2019). Effects of innovation in automated vehicles on occupant compartment designs, evaluation, and Safety: A review of public marketing, literature, and standards. *SAE Technical Papers, 2019-April*(April). <https://doi.org/10.4271/2019-01-1223>
- Hu, T., Jha, S., & Busso, C. (2020). Robust Driver Head Pose Estimation in Naturalistic Conditions from Point-Cloud Data. *IEEE Intelligent Vehicles Symposium, Proceedings*, 1176–1182.
<https://doi.org/10.1109/IV47402.2020.9304592>
- Lee, S. C., Yoon, S. H., & Ji, Y. G. (2021). Effects of Non-Driving-Related Task Attributes on Takeover Quality in Automated Vehicles. *International Journal of Human-Computer Interaction*, 37(3), 211–219.
<https://doi.org/10.1080/10447318.2020.1815361>
- Lu, Z., Happee, R., Cabrall, C. D. D., Kyriakidis, M., & de Winter, J. C. F. (2016). Human factors of transitions in automated driving: A general framework and literature survey. *Transportation Research Part F: Traffic Psychology and Behaviour*, 43(November), 183–198. <https://doi.org/10.1016/j.trf.2016.10.007>
- Martinez-Gonzalez, A. N., Villamizar, M., Canevet, O., & Odobez, J. M. (2020). Efficient Convolutional Neural Networks for Depth-Based Multi-Person Pose Estimation. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(11), 4207–4221.
<https://doi.org/10.1109/TCSVT.2019.2952779>
- Monnier, G., & Wang, X. (2009). A Motion Simulation Tool for Automotive Interior Design. In Vincent G. Duffy (Ed.), *Handbook of digital human modeling: research for applied ergonomics and human factors engineering*. CRC Press Taylor and Francis Group.
- Naujoks, F., Befelein, D., Wiedemann, K., & Neukum, A. (2018). A review of non-driving-related tasks used in studies on automated driving. *Advances in Intelligent Systems and Computing*, 597(June), 525–537.
https://doi.org/10.1007/978-3-319-60441-1_52
- Née, M., Contrand, B., Orriols, L., Gil-Jardiné, C., Galéra, C., & Lagarde, E. (2019). Road safety and distraction, results from a responsibility case-control study among a sample of road users interviewed at the emergency room. *Accident Analysis and Prevention*, 122(October 2018), 19–24.
<https://doi.org/10.1016/j.aap.2018.09.032>
- Plantard, P., Hubert, H. P., & Multon, F. (2017). Filtered pose graph for efficient kinect pose reconstruction. *Multimedia Tools and Applications*, 76(3), 4291–4312. <https://doi.org/10.1007/s11042-016-3546-4>
- Roth, M., & Gavrilu, D. M. (2019). DD-pose - A large-scale driver head pose benchmark. *IEEE Intelligent Vehicles Symposium, Proceedings, 2019-June*, 927–934.
<https://doi.org/10.1109/IVS.2019.8814103>
- Shotton, J. (2011). Real-time human pose recognition in parts from single depth images. *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 411, 119–135.
<https://doi.org/10.1109/CVPR.2011.5995316>
- Shotton, J., Girshick, R., Fitzgibbon, A., Sharp, T., Cook, M., Finocchio, M., Moore, R., Kohli, P., Criminisi, A., Kipman, A., & Blake, A. (2013). Efficient Human Pose Estimation from Single Depth Images. *Pattern Analysis and Machine Intelligence, IEEE Transactions On*, 35(12), 2821–2840.
<https://doi.org/10.1109/TPAMI.2012.241>
- Torres, H. R., Oliveira, B., Fonseca, J., Queirós, S., Borges, J., Rodrigues, N., Coelho, V., Pallauf, J., Brito, J., & Mendes, J. (2019). Real-Time Human Body Pose Estimation for In-Car Depth Images. *IFIP Advances in Information and Communication Technology*, 553, 169–182.
https://doi.org/10.1007/978-3-030-17771-3_14
- Venturelli, M., Borghi, G., Vezzani, R., & Cucchiara, R. (2017). From depth data to head pose estimation: A Siamese approach. *VISIGRAPP 2017 - Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 5, 194–201.
<https://doi.org/10.5220/0006104501940201>
- Wang, H. Y., Zhao, M. M., Beurier, G., & Wang, X. G. (2019). Automobile Driver Posture Monitoring Systems: A Review. *Zhongguo Gonglu Xuebao/China Journal of Highway and Transport*, 32(2), 1–18.
- Xing, Y., Lv, C., Zhang, Z., Wang, H., Na, X., Cao, D., Velenis, E., & Wang, F. Y. (2017). Identification and Analysis of Driver Postures for In-Vehicle Driving Activities and Secondary Tasks Recognition. *IEEE Transactions on Computational Social Systems*, 5(1), 95–108. <https://doi.org/10.1109/TCSS.2017.2766884>
- Yoon, S. H., & Ji, Y. G. (2019). Non-driving-related tasks, workload, and takeover performance in highly automated driving contexts. *Transportation Research Part F: Traffic Psychology and Behaviour*, 60, 620–631. <https://doi.org/10.1016/j.trf.2018.11.015>
- Yuen, K., & Trivedi, M. M. (2018). Looking at hands in autonomous vehicles: A convnet approach using part affinity fields. *ArXiv*, 5(3), 361–371.
<https://doi.org/10.1109/tiv.2019.2955369>
- Zhang, Z. (2002). A Flexible New Technique for Camera Calibration (Technical Report). *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), 1330–1334. <https://doi.org/10.1109/34.888718>
- Zhao, M., Beurier, G., Wang, H., & Wang, X. (2021a). Driver posture monitoring in highly automated vehicles using pressure measurement. *Traffic Injury Prevention*, 0(0), 1–6.
<https://doi.org/10.1080/15389588.2021.1892087>
- Zhao, M., Beurier, G., Wang, H., & Wang, X. (2021b). Exploration of Driver Posture Monitoring Using Pressure Sensors with Lower Resolution. *Sensors*, 1–13.
- Zhao, M., Georges, B., Wang, H., & Wang, X. (2020). A Pipeline for creating in-vehicle posture database for developing driver posture monitoring systems. *DHM2020*, 0, 187–196.
<https://doi.org/10.3233/ATDE200024>

Analysing Driver (In)Attentiveness: Towards a Cognitive Complexity Model Combining Visuospatial and Interactional Parameters

Vasiliki Kondyli^{1,*}, Mehul Bhatt¹

¹Örebro University, Sweden

Email: *vasiliki.kondyli@oru.se

We investigate the role of visuospatial environmental cues on driver (in)attention in everyday naturalistic driving situations. We develop a cognitive model of visuospatial complexity incorporating two critical aspects influencing visual (in)attention: (1) multimodal interaction mechanisms such as gesture, joint attention amongst roadside stakeholders (e.g. pedestrians, cyclists, drivers); and (2) visuospatial environmental features such as clutter, motion, environmental structure.

Our research emphasises the manner in which a cognitive human-factors guided model to analyse attentiveness can be applied to systematically explore the effects of a combination of environmental and interactional characteristics on visual attention in naturalistic driving. We position the application of the developed cognitive model to serve a foundational purpose in the training and testing of novel driver assistance technologies, e.g., from the viewpoint of systematic compliance with human-centered design guidelines.

1 The Complexity of Naturalistic Driving

Everyday driving involves a number of complex perceptual and cognitive tasks that contribute to successful navigation in the streetscape, cruising, overtaking, as well as in safe and efficient interactions with other street stakeholders –such as pedestrians, cyclists, drivers– in intersections of crossings. Human abilities to acquire relevant information from the dynamic environment, detect changes, track multiple objects, etc., which are necessary for coping with a number of challenges during driving, are affected by a plethora of environmental factors (among other aspects). For instance, rapid and frequent environmental changes in traffic demand greater attention and more timely responses than a rural one way street (Beanland, Filtness, and Jeans 2017). The range of environmental aspects involved in the dynamic driving environment can be described through the notion of environmental complexity, or by accounting for human perceptual and cognitive abilities, we use the notion of visuospatial complexity.

The effect of visual complexity on human perception has been studied in many areas, including cognitive science, psychology, human-computer interaction (e.g., Braun et al. 2013; Pieters, Wedel, and Batra 2010; Tuch et al. 2009). However, the notion of visual complexity remains elusive and several definitions have been proposed, originally focusing on static images and later also referring to dynamic stimuli. Visual complexity has been broadly de-

finied as the level of detail and intricacy contained within an image or a scene (Snodgrass and Vanderwart 1980). Adjusted to the case of real-world dynamic stimuli for everyday interactions with the environment, we consider the extended term of visuospatial complexity as the combination of visual and spatial characteristics that both coexist in dynamic naturalistic environment where a person acts. Perceiving a visuospatial stimulus as more or less complex has been suggested to be influenced by several factors, including type and quantity of elements contained, their spatial distribution or layout, variety of colors, etc. (Palumbo et al. 2014).

Driving performance has been shown to depend on visuospatial complexity, but also on the complexity of the driving task, and other factors that affect cognitive resources, such as individual differences, fatigue, age, dual task requests Doyon-Poulin, Ouellette, and Robert 2012; Smith and Evans 2013. Furthermore, previous studies on interpersonal communication examining the relation between multimodal interaction and cognitive load suggests that people tend to utilise multimodal signals when cognitive load increases due to task difficulty, communication or environmental complexity Oviatt, Coulston, and Lunstord 2004.

2 Driver (In)Attentiveness: Interactional and Environmental Factors

Visuospatial complexity is typically regarded to have a negative effect on the perceptual system, as it is mostly connected to “sensory overload”

| COGNITIVE ATTRIBUTES | Parameters |
|--|--|
| A. Embodied Multimodal Interactions | |
| MODALITIES | Speech - Head Movements - Facial Expressions - Gestures - Body Postures - Gaze - Auditory Cues |
| MODE | Explicit Interaction - Implicit Interaction |
| METHOD | Formal Device-based - Informal Device-based - Formal Body-based - Informal Body-based |
| LEVELS OF SOCIAL ATTENTION | Individual - Monitoring - Common - Mutual - Joint (Shared) |
| B. Visuospatial Complexity | |
| QUANTITATIVE ATTRIBUTES | Size: Width - Depth - Height Clutter: Variety of Colors - Variety of Shapes and Objects - Objects Density - Edges Density - Luminance - Saliency - Target-background similarity |
| STRUCTURAL ATTRIBUTES | Repetition - Symmetry - Order - Homo-Heterogeneity - Regularity - Openness - Grouping |
| DYNAMIC ATTRIBUTES | Motion - Flicker - Speed - Direction |
| Complexity of Human Performance | |
| PERFORMANCE EVALUATION | Search time - Accuracy |
| PHYSIOLOGICAL MEASUREMENTS | Eye-tracking: Latency of first saccade - Total number of fixations - Number of fixations (no distractors) - Scanpath ration - Mean fixation duration |

Figure 1: The taxonomy of cognitive attributes constituting visuospatial complexity and the metrics of Human Performance.

that may impair cognitive function. There exists suggestive evidence that complexity impairs perceptual sensitivity, as for example in the case of multiple object tracking, where every additional moving object diminishes a participant’s ability to track objects (Pylyshyn and Storm 1988). However, increased complexity might improve perceptual sensitivity under some task demands. Recent findings suggest that despite increasing informational load, complexity can serve to ground and facilitate perceptual sensitivity (Ellis and Turk-Browne 2019).

To provide a holistic approach of visuospatial complexity and its effects on visual (in)attention during driving, we employ an empirical basis for our complexity model. We analyse 25 real-world driving scenes, with 75 interaction scenarios with street stakeholders (e.g. crossing with gestures, overhead checking and overtaking, joint attention by gaze or gesture initiations) from 25 different locations worldwide (e.g., Australia, South Korea, India, USA, UK, China). Systematic qualitative and quantitative analysis of this set of dynamic driving stimuli included semantic annotations of environmental cues as well as the multimodal interactions of interpersonal communication in the street. This analysis also led to select instances that we replicated in a virtual environment (VR) such that the instances overall encompass an entire spectrum of complexity levels. A series of behavioural studies are then employed for the evaluation of a visuospatial complexity scale based on human experience and perceptual performance. Overall, the model involves two sets of cognitive parameters (A & B; Fig. 1):

A. Multimodal Interactions

To investigate aspects of visual attention during driving, we need to consider the nature of multimodal interactions between humans during everyday events in the streetscape. Multimodal interactions highly vary and they can convey very different meanings depending on the users involved (e.g. pedestrian, cyclist, driver), their intentions, and activities in the streetscape (e.g. stop, accelerate, turn, cross), as well as the environmental and situational context (e.g. scene complexity, demographics, culture). In essence, multimodal interactions correspond to the characteristics of the interpersonal communication between roadside stakeholders focusing on the combination of modalities involved, the mode and method of delivering the message as well as the social attention achieved between the parts (Kondyli and Bhatt 2020).

B. Visuospatial Characteristics

We examine the range of complexity for visuospatial stimuli based on the combination of three categories of attributes pertaining to environmental characteristics: *quantitative*, *structural* and *dynamic*. Quantitative aspects refer to the size of the space, the quantity and quality of the objects involved. Structural examines how the objects are positioned in space (e.g. order, heterogeneity), while dynamic refer to motion and various directions of moving objects (Kondyli, Bhatt, and Suchan 2020). A systematic analysis of different combinations of attributes can provide a better understanding of the aggravation or counterbalance dynamics between the attributes and their effect on human behaviour. To empirically define a model of visuospatial complexity for naturalistic envi-

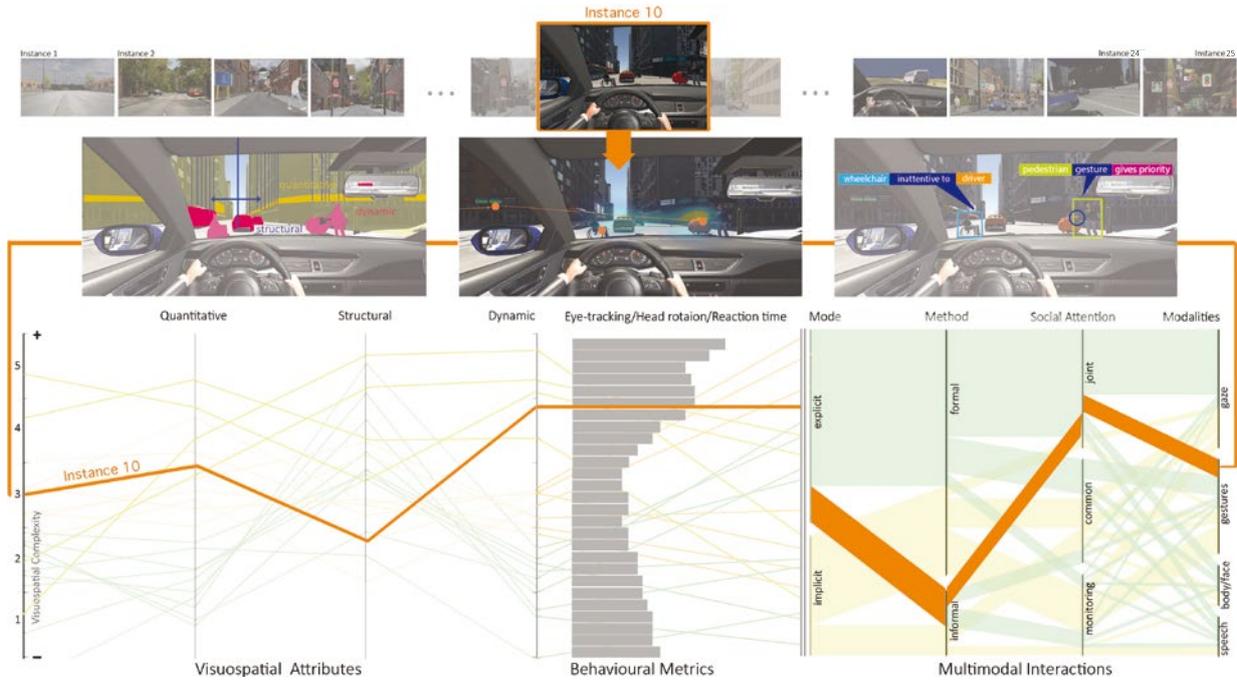


Figure 2: A conceptual illustration of the distribution of stimuli complexity in a dataset focusing on safety-critical situations. Diverging complexity levels are analysed vis-a-vis driver performance parameters such as those in Fig. 1.

ronments, we use the taxonomy introduced in Fig. 1) to develop a number of virtual environments that differ on the combination of attributes involved, as well as on the degree of each attribute the environment contains.

3 Complexity of Human Performance

Driver (in)attention can be analysed vis-a-vis the interactional and environmental characteristics in (A) and (B) respectively, i.e., via correlation with driver’s behaviour, perception, reasoning and decision making as captured via behavioural and psychophysical measures during embodied active driving tasks (in the real world, or in VR) (Fig. 1). The evidence-based performance parameters refer to a combination of behavioural (e.g. detection rate, reaction time) and physiological measurements (e.g. eye-tracking, head rotation, steering, braking, acceleration) of subjects during driving. In a series of behavioural studies we investigate the effect of environmental attributes, and multimodal interactions in subjects’ performance through naturalistic driving tasks focused on visual attention, such as visual search task, cognitive load or predictions during active driving. The results of the studies inform the cognitive model about the combinations of aspects that promote or aggravate performance, and further

help identifying the scale of complexity based on empirical evidence under naturalistic conditions. For instance, preliminary results on visual search during a driving task suggest that structural cues can counterbalance the effect of extensive clutter and limit the number of fixations indicating a medium level of complexity.

4 Human-centred Evaluation of Driving Assistance Technologies

In the context of autonomous driving research and the urban environment, evaluating a driving dataset (and AI system) with respect to human-centred factors involves an analysis of how well the range of different types of multimodal interactions as well as the levels of visuospatial complexity can be successfully handled or are represented in the instances of the dataset (Fig. 2).

In addition to serving its crucial purpose as an analytical tool for studying driver attentiveness, our empirically based model of visuospatial complexity can be used as a basis for systematic analysis of various aspects of visual attention during driving under naturalistic conditions. From the viewpoint of characterising the multi-faceted nature and complexity of everyday (driving) situations, the proposed cognitive model promises to centralise human factors as a crucial aspect for the design, evaluation, and

deployment of human-centred visual sensemaking technologies within autonomous driving systems. (Suchan, Bhatt, and Varadarajan 2021). Moreover, the model by including a range of human-factors can provide a common reference frame and guidelines for the training and testing of datasets for autonomous systems that focus on human-human and human-machine interactions. Furthermore, datasets that follow the cognitive model of complexity can also be part of a common platform for shared stimuli of naturalistic driving environment and hence to facilitate reproducibility in experimental work and in empirical studies under ecological valid conditions.

References

- Beanland, V., Filtness, A. J., and Jeans, R. (2017). “Change detection in urban and rural driving scenes: Effects of target type and safety relevance on change blindness”. In: *Accident Analysis Prevention* 100, pp. 111–122. ISSN: 0001-4575. DOI: <https://doi.org/10.1016/j.aap.2017.01.011>. URL: <https://www.sciencedirect.com/science/article/pii/S0001457517300362>.
- Braun, J., Amirshahi, S. A., Denzler, J., and Redies, C. (2013). “Statistical image properties of print advertisements, visual artworks and images of architecture”. In: *Frontiers in Psychology* 4. ISSN: 1664-1078. DOI: [10.3389/fpsyg.2013.00808](https://doi.org/10.3389/fpsyg.2013.00808). URL: <https://www.frontiersin.org/article/10.3389/fpsyg.2013.00808>.
- Pieters, R., Wedel, M., and Batra, R. (2010). “The Stopping Power of Advertising: Measures and Effects of Visual Complexity”. In: *Journal of Marketing* 74.5, pp. 48–60. ISSN: 0022-2429. URL: <http://www.jstor.org/stable/41228573> (visited on 04/20/2022).
- Tuch, A., Bargas-Avila, J., Opwis, K., and Wilhelm, F. (Sept. 2009). “Visual complexity of websites: Effects on users’ experience, physiology, performance, and memory”. In: *International Journal of Human-Computer Studies* 67, pp. 703–715. DOI: [10.1016/j.ijhcs.2009.04.002](https://doi.org/10.1016/j.ijhcs.2009.04.002).
- Snodgrass, J. and Vanderwart, M. (1980). “A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity.” In: *Journal of Experimental Psychology: Human Perception and Performance* 6, pp. 174–215.
- Palumbo, L., Ogden, R., Makin, D., and Bertamini, M. (2014). “Examining visual complexity and its influence on perceived duration”. In: *Journal of Vision* 14.3, pp. 1–18.
- Doyon-Poulin, F., Ouellette, B., and Robert, J. (2012). “Review of visual clutter and its effects on pilot performance: New look at past research”. In: *Digital Avionics Systems Conference (DASC)*. IEEE/AIAA 31st.
- Smith, C. and Evans, C. (2013). “A new heuristic for capturing the complexity of multimodal signals”. In: *Behavioral Ecology and Sociobiology* 67.9.
- Oviatt, S., Coulston, R., and Lunstord, R. (2004). “When Do We Interact Multimodally? Cognitive Load and Multimodal Communication Patterns”. In: ICMI. Pennsylvania, USA.
- Pylyshyn, Z. W. and Storm, R. (1988). “Tracking multiple independent targets: evidence for a parallel tracking mechanism.” In: *Spatial vision* 3 3, pp. 179–97.
- Ellis, T. C. and Turk-Browne, B. N. (2019). “Complexity can facilitate visual and auditory perception”. In: *J Exp Psychol Hum Percept Perform* 45.9, pp. 1271–1284.
- Kondyli, V. and Bhatt, M. (2020). “Multimodality on the Road: Towards Evidence-Based Cognitive Modelling of Human Interactions in Everyday Roadside Situations”. In: *6th International Digital Human Modeling Symposium 2020 (DHM2020)*. Skövde, Sweden.
- Kondyli, V., Bhatt, M., and Suchan, J. (2020). “Multimodal Interaction in Autonomous Driving. Towards Human Visual Perception Driven Standardisation and Benchmarking”. In: *STAIRS @ ECAI 2020: 9th European Starting AI Researchers Symposium (STAIRS)., at ECAI 2020, the 24th European Conference on Artificial Intelligence (ECAI)*.
- Suchan, J., Bhatt, M., and Varadarajan, S. (2021). “Commonsense visual sensemaking for autonomous driving - On generalised neurosymbolic online abduction integrating vision and semantics”. In: *Artif. Intell.* 299, p. 103522. DOI: [10.1016/j.artint.2021.103522](https://doi.org/10.1016/j.artint.2021.103522). URL: <https://doi.org/10.1016/j.artint.2021.103522>.

Train driver attention is influenced by the type of railway signalling system

Katja Kircher^{1*}, Tomas Rosberg¹, Birgitta Thorslund¹, Christer Ahlström^{1,2}, Erik Prytz³, Lisa Bernheim³, Niklas Olsson¹, Peter Moertl⁴

¹ VTI, Linköping, Sweden

² Department of Biomedical Engineering, Linköping University, Linköping, Sweden

³ Department of Computer and Information Science, Linköping University, Linköping, Sweden

⁴ Virtual Vehicle Research GmbH, Graz, Austria

*katja.kircher@vti.se

Abstract: The European Rail Traffic Management System (ERTMS) will replace national standards with the aim to promote cross-border traffic and enhance efficiency. The transition involves a shift from lineside signalling to mostly in-cabin information via a Driver Machine Interface (DMI). Previous research indicates that this may lead to a decrease in driver attention to the outside world and to a decrease in workload, leading to boredom. Using a train simulator, 41 participants drove the same track with the ERTMS system and the Swedish national standard (ATC) while wearing eye-tracking equipment. Subjective workload and boredom assessments were made after each drive. An analysis of the first set of reduced data (15 participants) showed that the formal attentional requirements like the monitoring of speed changes and signals were fulfilled in almost all cases, regardless of system. Overall, however, the data indicate that in line with previous research the drivers focus their attention more to the inside of the train when using the ERTMS system. This is corroborated by the finding that horn blowing is slightly delayed with the ERTMS system. Perceived workload was generally low, with the ERTMS system experienced to be more boring. We draw the preliminary conclusion that while formal attentional requirements are fulfilled for both systems, the ERTMS system likely has a tendency to pull the drivers' overall attention inwards. Given that for the ERTMS system most relevant information is presented inside of the train on the DMI, this is not surprising, but needs to be addressed by the authorities.

1. Introduction

The European Rail Traffic Management System (ERTMS) is a set of standards for management and interoperation of signalling for railways by the European Union. ERTMS aims to promote cross-border traffic and enhance efficiency of train transports by replacing national signalling equipment and operational procedures with a European standard for train control and command systems. The transition to the new ERTMS system involves a shift from lineside signalling to mostly in-cabin information provision via a Driver Machine Interface (DMI). Presenting most information via the in-cabin DMI instead of via lineside signalling will inevitably change the drivers' visual information sampling behaviour. With ERTMS, events outside the cabin have been found to be attended less (Buksh, Sharples, Wilson, Morrisroe, & Ryan, 2013; Naghiyev, Sharples, Carey, Coplestone, & Ryan, 2014; van der Weide, De Bruijn, & Zeilstra, 2017). For example, van der Weide et al. (2017) found that the share of outside glances was reduced from 75% for lineside signalling to 40% for ERTMS. A clear DMI and increased automation can reduce perceived workload, whereas complex speed profiles and badly planned speed changes may cause problems. Decreased workload, to the point that it leads to boredom due to underload, has also been put forward as a potential risk.

In this study we investigate two different but related questions. Firstly, we assess if train drivers fulfil the formal criteria for attentive driving, that is, whether drivers sample the necessary information on speed changes, signals, and other relevant information, both when using ERTMS and when using a lineside signalling system (ATC). To this end,

we employ the theory of Minimum Required Attention (MiRA, Kircher & Ahlstrom, 2017), which provides a framework for identifying formal attentional requirements and assessing if these requirements are met. Secondly, we investigate how the shift from ATC to ERTMS influences the drivers' propensity to monitor the outside scene. Subjective workload data complement the findings.

2. Method

A train simulator study was conducted using a within-subjects design with one ATC and three ERTMS conditions in counterbalanced order. The analogous ATC system provided speed information, signals and signage externally. Speed information could also be obtained internally via the dashboard. The digital ERTMS provided most information internally via the DMI (except external whistle boards, indicating the obligation to blow the horn), with external signage and auditory information indicating where new internal information is given. Here, the ATC condition is compared with an ERTMS condition providing the same speed profile.

2.1 Participants

All 41 participants (11 female; mean age 41 years) were train drivers familiar with both ATC and ERTMS systems. On average the participants had 6.9 years of experience from train driving and 4.8 years of driving with ERTMS. Most of the participants (34) were drivers of passenger trains, five were freight train drivers, two drove work train. As compensation the participants received 500 SEK.

2.2 Equipment and route

The simulator was a fixed-base mock-up of a Bombardier Regina EMU (X55) train cockpit (Thorslund, Rosberg, Lindström, & Peters, 2019). The route was a simulation of the main track at Ådal line, specifically the stretch between Ramvik and Dynäs containing a stop in the town of Kramfors. The route was 19 km long and driving one condition took approximately 15 minutes. The participants' gaze direction was tracked with Pupil Labs Invisible glasses (Pupil Labs, Berlin, Germany).

2.3 Procedure

The participants were given the chance to familiarise themselves with the simulator. They were then equipped with the eye tracker. The participants were instructed to drive at a normal pace without being overly cautious. After each condition the participants reported subjective workload on the NASA-RTLX (Hart, 2006) by rating their effort from very low to very high and by answering selected items of the Multidimensional State Boredom Scale (MSBS) with ratings from strongly disagree to strongly agree (Fahlman, Mercer-Lynn, Flora, & Eastwood, 2013).

2.4 Analysis

Data from a 10 km long section were chosen for the analysis. This section contained three whistle boards requiring the driver to sound the horn (external information in both conditions), several speed changes, signals of different types and one stop. For each of these items, a zone within which the corresponding information had to be sampled was identified. For example, for the whistle board this zone started 100 m in front of the board and ended at the board. It was also identified how the relevant information could be obtained. For the whistle board, this was via glancing at the board, for speed information this could mean glancing at the speed sign (ATC) or the information on the DMI (ATC and ERTMS).

For each driver, we analysed whether the required information was sampled and from which source it was sampled. For the time it took passing two stationary oncoming trains and in the area around level crossings ("critical phases") gaze behaviour was coded glance by glance, indicating the glance target.

3. Results

The results on visual sampling in this abstract are preliminary and based on 15 participants. A complete analysis will be available at the time of the conference. The results for sounding the horn are for all participants.

Even though the speed profile and the route were the same for both conditions, the number of attention requirements differed, because the convention how information is provided differs between systems. Per participant the number of requirements on the analysed track section were (ATC/ERTMS): whistle board 3/3; U-sign 1/1; signal, any type 20/11; speed information 10/7.

3.1 Visual sampling

It was very uncommon for the drivers to miss visual sampling of relevant information as identified by the MiRA-

theory – this happened in only 11 cases out of 840 (4 ATC; 7 ERTMS). In an additional 38 cases the required information was not sampled foveally, but probably with peripheral vision (29 ATC; 9 ERTMS).

The glance distribution between the inside and outside of the train was analysed for the critical phases. With the ATC system, drivers spent 26 % of the time during a critical phase looking at the dashboard. For ERTMS this was 44 %. Glances towards the dashboard were longer on average for ERTMS (0.94 s vs. 0.76 for ATC), and they were 1.7 times as many. Both the maximum duration for a single glance to the dashboard and the 85th percentile were larger for ERTMS (max: 5.5 s for ERTMS, 3.9 for ATC; 85th: 1.4 s for ERTMS, 1.24 s for ATC). Significance testing will be done when all data are reduced.

3.2 Sounding the horn

Based on data from 45 runs, the horn was not blown at the whistle board on two occasions (one ATC, one ERTMS). Overall, with ATC the horn was blown for the first time before the board was passed in 54% of the cases, with ERTMS this happened in 45% of the cases. For the remaining cases the horn was blown after the board was passed (except the two times where it was not blown at all).

3.3 Workload/Mental load

The driving task was perceived as easy, as shown by low workload ratings in both conditions. The average RAW-TLX ratings were never higher than 34.8 on the scale from 0 (low effort) to 100 (high effort). On average, boredom was rated on the lower part of the scale for all questions except for "It was all repetitive and routine for me", which had a mean value close to 5 (somewhat agree) on the scale from 1 (strongly disagree) to 7 (strongly agree). According to one of the items in the MSBS, drivers felt more bored when driving with ERTMS, $t(40) = 2.06$, $p = 0.046$.

4. Discussion

The preliminary results indicate that while the formal attentional requirements as operationalised by the MiRA theory are fulfilled in almost all cases for both systems, the glance distribution in critical phases indicates that the ERTMS system likely has a tendency to pull the drivers' overall attention inwards. This is corroborated by the operation of the horn, which shows a similar pattern. While the drivers rarely miss blowing the horn, they tend to act later in the ERTMS condition.

Given the location where most relevant information pertaining to driving the train is presented, this shift does not come as a surprise. If information is expected to be presented on the DMI, this is the natural place to monitor. However, unexpected and unpredictable events, especially critical ones, are more likely to occur outside.

5. Conclusions

By design, the ERTMS system requires drivers to sample necessary information from the DMI, which has the concerning consequence that drivers are less likely to monitor the outside environment also in critical phases. This needs to be investigated further and addressed by system developers.

6. Acknowledgments

This research was funded by The Virtual Vehicle Research GmbH, InfraSweden 2030 - Solutions for sustainable transport infrastructure, and The Swedish Transport Administration. We would like to thank all participants for contributing with their time and insights.

References

- Buksh, A., Sharples, S., Wilson, J., Morrisroe, G., & Ryan, B. (2013, 15-18 April 2013). Train automation and control technology – ERTMS from users' perspectives. Paper presented at the Contemporary Ergonomics and Human Factors, Cambridge, UK.
- Fahlman, S. A., Mercer-Lynn, K. B., Flora, D. B., & Eastwood, J. D. (2013). Development and Validation of the Multidimensional State Boredom Scale. *Assessment*, 20(1), 68-85.
doi:10.1177/1073191111421303
- Kircher, K., & Ahlstrom, C. (2017). Minimum Required Attention: A human-centered approach to driver inattention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 59(3), 471-484.
doi:10.1177/0018720816672756
- Naghiyev, A., Sharples, S., Carey, M., Coplestone, A., & Ryan, B. (2014, 7–10 April 2014). ERTMS train driving - incab vs. outside: an explorative eye-tracking field study. Paper presented at the Contemporary Ergonomics and Human Factors, Southampton, UK.
- Thorslund, B., Rosberg, T., Lindström, A., & Peters, B. (2019, June 17-20, 2019). User-centered development of a train driving simulator for education and training. Paper presented at the 8th International Conference on Railway Operations Modelling and Analysis (ICROMA 2019), Norrköping, Sweden.
- van der Weide, R., De Bruijn, D., & Zeilstra, M. (2017). ERTMS pilot in the Netherlands – impact on the train driver. Paper presented at the International Human Factors Rail Conference, London, UK.

The influence of alcohol and automation on drivers' visual behavior during test track driving

Emma Tivesten^{1,2*}, Viktor Broo¹, Mikael Ljung Aust¹

¹ Volvo Cars Safety Centre, PV22, Göteborg, Sweden

² Chalmers University of Technology, Göteborg, Sweden

*emma.tivesten@volvocars.com

Abstract: Background: Driving under the influence of alcohol increases crash risk and is a major contributing factor to severe crashes. **Method:** A test-track study investigated the effects of alcohol intoxication on drivers' visual behavior both when just driving and when engaged in secondary tasks. Twenty-six participants performed two drives: 1) sober baseline, 2) with a target Blood Alcohol Concentration at 0.1%. The participants drove in either manual, assisted, or autonomous drive (AD) mode. **Results:** Intoxication influenced glance behavior in all driving modes. It was most evident during visually demanding secondary tasks where it resulted in longer single and total off-path glance durations. Additionally, in manual mode, more than one out of four of the drivers displayed gaze concentration to the forward roadway when intoxicated. For sober driving, the difference in off-path glance durations between manual and assisted mode were moderate, while there was a huge shift towards long off-path glance durations in AD mode. This mode effect was further amplified by intoxication. **Conclusions:** Intoxication clearly effects drivers' eyes on/off road glance behavior and can therefore be viewed as a promising indicator of alcohol intoxication. However, relatively refined metrics that account for both driving mode and secondary task engagement will be required to reliably distinguish sober from drunk driving. Still, driver monitoring systems that can measure eye movements in real-time can be used to detect driver impairment, and consequently be used for in-vehicle countermeasures.

1. Introduction

Drunk driving is a major contributing factor to fatal crashes (SNRA, 2020; NHTSA, 2017; WHO; 2007). Also, there is a well-established dose-dependent link between Blood Alcohol Concentration (BAC) and crash risk, that increases exponentially for BAC greater than 0.1% (Blomberg et al., 2009).

Alcohol impairs skills necessary for safe driving (Garrisson et al, 2021; Ogden and Moskowitz, 2004; Martin et al, 2013), and affect driving performance metrics related to lane-keeping (Martin et al., 2013; Jongen et al, 2018). While visually demanding secondary tasks alone reduce driving performance (Irwing et al, 2015), the combined effect of alcohol and secondary tasks interact to further impair driving performance (Harrisson & Fillmore, 2011; Rakauskas et al, 2008; Van Dyke and Fillmore, 2015).

A new challenge is that performance degradation in lateral control can no longer provide viable impairment detection when driving becomes assisted, and the vehicle controls lateral position. Consequently, additional metrics to recognize driver impairment are needed.

Here, a promising candidate is Driver Monitoring Systems (DMS) that include eye tracking. Alcohol intoxication leads to increased gaze concentration to the road center, fewer fixations to the peripheral areas, and longer time to read route signs (Belt, 1979; Moskowitz and Robinson, 1988; Moskowitz and Ziedman, 1979). Lee et al. (2010) suggested that metrics that cumulate over time, including gaze concentration to the forward roadway, may be a suitable metrics to detect alcohol impairment.

In the current study, we investigated drivers' visual behavior during a sober baseline drive and compared it to a second intoxicated drive for three driver groups that were

assigned to different driving modes (manual, assisted, autonomous drive (AD) mode) including segments with and without secondary tasks.

2. Method

2.1 Participants

The participants (N=26; 17 male and 9 female; age: 25-66 years, M=42.9, SD=12.5) were all moderate drinkers. They were divided into three groups assigned to manual (N=10), assisted (N=8), or AD mode (N=8).

2.2 Test environment and equipment

The study was performed in a rural road environment on a test-track. The test vehicle (TV) was a Volvo XC90, and speed was limited to 50 km/h. No other vehicles were present. A breathalyzer was used to estimate BAC. In assisted mode, test participants used the Pilot Assist function which performs both longitudinal and lateral control but requires supervision and hands on the wheel. AD mode was simulated by using the Pilot Assist function but with "hands-on-wheel"-reminders disabled. In AD mode, participants were allowed to disengage from driving but needed to remain available to take over if requested. A safety driver was present in the front passenger seat during all intoxicated drives and in AD mode.

2.3 Procedure

The participants practiced three visual-manual secondary tasks while seated in the stationary TV using the center stack display: 1) Tune the radio to a specific frequency, 2) Dial your own mobile phone number, 3) Set the in-car temperature.

The participants first performed a sober baseline drive. All participants drove manually the first 5 minutes, and then

in manual, assisted or AD mode for another 25 minutes. The secondary tasks were performed on straight road segments between 7 to 20 minutes into the drive. The baseline drive was followed by a drinking session to reach a target BAC of 0.1%. A second drive while intoxicated was then performed repeating the baseline drive procedure.

2.4 Data processing

Four video segments were extracted for manual annotations from each drive (Table 1). Gaze direction was coded as time series data at 20 Hz and merged into three categories: on-path, off-path or unknown.

Table 1 Selected segments including: instructed secondary tasks, the range of duration in seconds, and mode including Manual, Assisted, or AD mode.

| Segment | Task | Duration (s) | Mode |
|---------|-------|--------------|----------|
| S1 | No | 30 | M |
| S2 | Radio | 12 – 54 | M, A, AD |
| S3 | Dial. | 11 – 114 | M, A, AD |
| S4 | Temp. | 6 – 32 | M, A, AD |

2.5 Analysis

The metrics PRC (percent road center) and GF-off (off-path glance frequency) were investigated for all segments. Additionally, TGT (total glance time off-path), GD>2s (percentage of off-path glances longer than 2 seconds), and MaxGD (maximum off-path glance duration) were investigated for the three task segments S2-S4.

Wilcoxon signed rank test was used to compare each segment and metric during the intoxicated drive to the baseline. An accepted false discovery rate at 5% was applied to adjust for multiple testing (Benjamini and Hochberg, 1995).

3. Results

When just driving in manual mode, PRC increase when intoxicated. In segment S1 (manual, no task), there was a statistically significant effect of intoxication resulting in higher PRC and lower GF-off compared to baseline (Fig. 1, Table 2). Also, 27% of the drivers (n=7/26) showed gaze concentration to the roadway (PRC > 92%; Victor and Larsson, 2004) when intoxicated, while this was not present in baseline (n=0/26).

On the other hand, both the median PRC and GF-off were consistently lower during secondary tasks (S2-S4) in the intoxicated drives across all modes. During secondary tasks there was also a consistent trend of higher median for all off-path glance duration metrics (TGT, MaxGD, GD>2s) across all tasks and modes during intoxication. Fig. 2 illustrates the difference in off-path glance distributions between modes and drives.

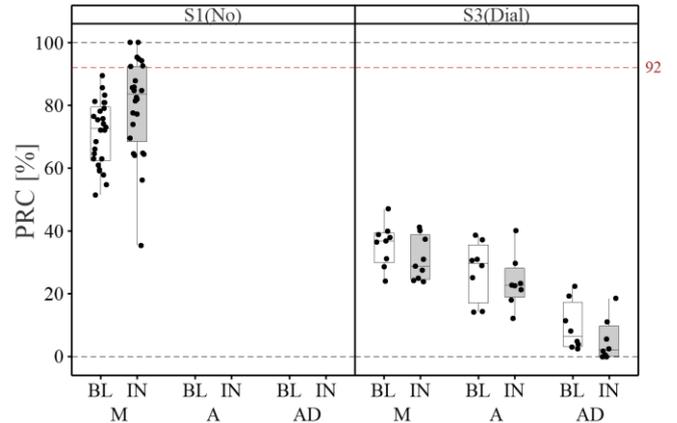


Fig. 1. Boxplots of PRC for segment S1 (left panel) with no secondary task and segment S3 (right panel) with the dialling task. Each panel shows BaseLine and INtoxicated drives, as well as driving modes (Manual, Assisted, AD).

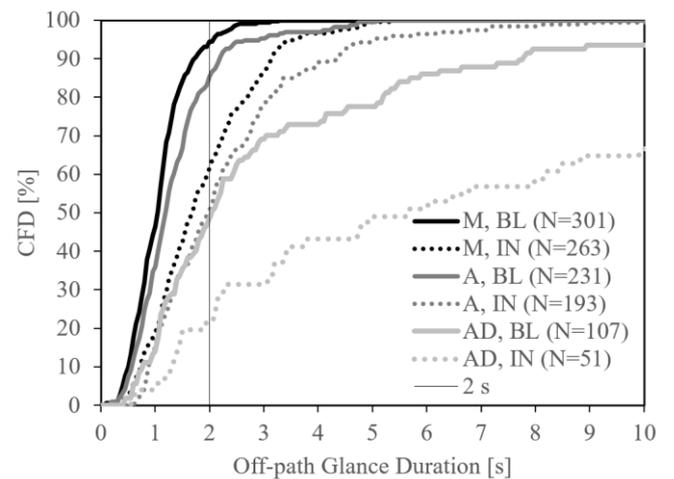


Fig. 2. Cumulative frequency distributions for all off-path glance durations during the three task segments combined. Solid lines represent BaseLine, dotted lines represent INtoxicated drives. The grey scale represents Manual, Assisted, and AD modes. The total number of off-path glances for each curve is presented in the legend.

The glance duration metrics were more sensitive in intoxication detection for the more visually demanding secondary tasks (i.e., Radio and Dialing, see Table 2). MaxGD was the most sensitive metric to detect intoxication during secondary tasks and consistently showed significant differences and high effect sizes between intoxication and baseline for all modes during the more demanding tasks (Table 2). TGT and GD>2s followed a similar pattern but with lower effect sizes.

Table 2 Overview of the results showing the comparisons that were statistically significant at 0.05 (*), 0.01 (**), and 0.001 (***) level, and indicating the effect sizes that were either moderate (Δ , $r = 0.31-0.49$) or high ($\Delta\Delta$, $r \geq 0.50$).

| Metric | Task | Manual | Assisted | AD |
|------------|------|--------|--------------------|-------------------------------------|
| PRC [%] | S1 | - | ** Δ | |
| | S2 | Radio | Δ | $\Delta\Delta$ ** $\Delta\Delta$ |
| | S3 | Dial. | Δ | Δ ** $\Delta\Delta$ |
| | S4 | Temp. | | |
| GF-off [N] | S1 | - | *** $\Delta\Delta$ | |
| | S2 | Radio | | Δ * Δ |
| | S3 | Dial. | | Δ * Δ |
| | S4 | Temp. | | |
| TGT [s] | S2 | Radio | ** $\Delta\Delta$ | * $\Delta\Delta$ $\Delta\Delta$ |
| | S3 | Dial. | * $\Delta\Delta$ | Δ ** $\Delta\Delta$ |
| | S4 | Temp. | Δ | Δ |
| GD>2s [%] | S2 | Radio | ** $\Delta\Delta$ | * $\Delta\Delta$ ** $\Delta\Delta$ |
| | S3 | Dial. | ** $\Delta\Delta$ | $\Delta\Delta$ Δ |
| | S4 | Temp. | | * Δ |
| MaxGD [s] | S2 | Radio | ** $\Delta\Delta$ | ** $\Delta\Delta$ ** $\Delta\Delta$ |
| | S3 | Dial. | ** $\Delta\Delta$ | * $\Delta\Delta$ ** $\Delta\Delta$ |
| | S4 | Temp. | Δ | ** $\Delta\Delta$ Δ |

4. Conclusions

This study investigated the effect of alcohol intoxication on drivers' visual behavior with participants driving either in manual, assisted, or AD mode.

Drivers glance behavior was influenced by intoxication in all driving modes. Intoxication increased PRC in manual mode when not performing a specific secondary task, which resulted in gaze concentration to the road in more than one out of four participants. During secondary tasks, intoxication resulted in lower PRC, lower number of off-path glances, longer total glance time, and longer off-path glance durations. The effect of intoxication was most evident during visually demanding secondary task.

These findings suggest that drivers' eyes on/off-path glance behavior likely can be used to detect alcohol intoxication in different driving modes, but also that visual time-sharing must be accounted for to interpret the effects correctly.

5. Acknowledgments

The authors would like to thank Regina Johansson, Elin Meltzer, Katerina Pistola, Joel Johansson, and Emma Kroon for preparing and performing data collection and video annotations. This research was supported by the European Project L3Pilot, grant number 72305, and approved by the Swedish Ethical Review Authority (Dnr 2019-05395).

References

- Belt, B. (1969). *Driver eye movements as a function of low blood alcohol concentrations*. Technical report, Driving Research Laboratory, Department of Industrial Engineering, Ohio State University
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Blomberg, R. D., Peck, R. C., Moskowitz, H., Burns, M., & Fiorentino, D. (2009). The Long Beach/Fort Lauderdale relative risk study. *Journal of Safety Research*, 40(4), 285–292. <https://doi.org/10.1016/j.jsr.2009.07.002>
- Garrison, H., Scholey, A., Ogden, E., & Benson, S. (2021). The effects of alcohol intoxication on cognitive functions critical for driving: A systematic review. *Accident Analysis and Prevention*, 154(December 2020), 106052. <https://doi.org/10.1016/j.aap.2021.106052>
- Harrison, E. L. R., & Fillmore, M. T. (2011). Alcohol and distraction interact to impair driving performance. *Drug and Alcohol Dependence*, 117(1), 31–37. <https://doi.org/10.1016/j.drugalcdep.2011.01.002>
- Irwin, C., Monement, S., & Desbrow, B. (2015). The Influence of Drinking, Texting, and Eating on Simulated Driving Performance. *Traffic Injury Prevention*, 16(2), 116–123. <https://doi.org/10.1080/15389588.2014.920953>
- Jongen, S., van der Sluiszen, N. N. J. J. M., Brown, D., & Vuurman, E. F. P. M. (2018). Single- and dual-task performance during on-the-road driving at a low and moderate dose of alcohol: A comparison between young novice and more experienced drivers. *Human Psychopharmacology*, 33(3), 6–11. <https://doi.org/10.1002/hup.2661>
- Lee, J. D., Fiorentino, D., Reyes, M. L., Brown, T. L., Ahmad, O., Fell, J., Ward, N., & Dufour, R. (2010). *Assessing the feasibility of vehicle-based sensors to detect alcohol impairment*. Report No. DOT HS 811 358. Washington, DC: National Highway Traffic Safety Administration.
- Martin, T. L., Solbeck, P. A. M., Mayers, D. J., Langille, R. M., Buczek, Y., & Pelletier, M. R. (2013). A review of alcohol-impaired driving: The role of blood alcohol concentration and complexity of the driving task.

Journal of Forensic Sciences, 58(5), 1238–1250.
<https://doi.org/10.1111/1556-4029.12227>

Moskowitz, H., & Robinson, C. (1988). *Effects of low doses of alcohol on driving-related skills: a review of the evidence*. National Highway Traffic Safety Administration, Report No.: DOT HS 807 280.
<https://rosap.nhtsa.gov/view/dot/1461>

Moskowitz, H., & Ziedman, K. (1979). Eye Movements and Skills Performance Measures under Alcohol in a Driving Simulator. *Proceedings of the Human Factors Society Annual Meeting*, 23(1), 389–393.
<https://doi.org/10.1177/107118137902300198>

NHTSA. (2017). *Alcohol Impaired Driving. Traffic safety facts report DOT HS 812450*. Washington, DC: National Highway Traffic Safety Administration. <https://doi.org/10.32964/tj16.10>

Ogden, E. J. D., & Moskowitz, H. (2004). Effects of alcohol and other drugs on driver performance. *Traffic Injury Prevention*, 5(3), 185–198.
<https://doi.org/10.1080/15389580490465201>

Rakauskas, M. E., Ward, N. J., Boer, E. R., Bernat, E. M., Cadwallader, M., & Patrick, C. J. (2008). Combined effects of alcohol and distraction on driving performance. *Accident Analysis and Prevention*, 40(5), 1742–1749.
<https://doi.org/10.1016/j.aap.2008.06.009>

SNRA. (2020). *Fakta om alkohol och narkotika i trafiken*. Retrieved from the Swedish National Road Administration website on 2021-12-15.
<https://www.trafikverket.se/resa-och-trafik/trafiksakerhet/Din-sakerhet-pa-vagen/Rattfylleri/fakta-om-alkohol-och-narkotika-i-trafiken/>

Van Dyke, N. A., & Fillmore, M. T. (2015). Distraction produces over-additive increases in the degree to which alcohol impairs driving performance. *Psychopharmacology*, 232(23), 4277–4284.
<https://doi.org/10.1007/s00213-015-4055-2>.

Victor, T., & Larsson, P. (2004). *Method and arrangement for interpreting a subjects head and eye activity* (Patent No. International publication number PCT WO 2004/034905 A1).
<https://patents.justia.com/patent/7460940>

WHO. (2007). *Drinking and Driving: a road safety manual for decision-makers and practitioners*. Geneva, Global Road Safety Partnership.
https://www.grsproadsafety.org/wp-content/uploads/DrinkingDriving_English.pdf

How does night-time driving and engagement in a cognitive distraction task affect detection of peripheral targets?

İbrahim Öztürk^{1*}, Natasha Merat¹

¹ Institute for Transport Studies, University of Leeds, Leeds, United Kingdom

*: i.ozturk@leeds.ac.uk

Abstract: Driver distraction is known to be a potential risk factor for traffic safety. Previous studies have shown that increased cognitive load can affect many driving outcomes, and lab-based studies have commonly used the detection-response task (DRT) to quantify the level of cognitive load from in-vehicle systems during driving. The aim of the present study is to examine the effects of varying levels of cognitive distraction (two levels of n-back task) on performance in the detection-response task, in day- and night-time driving conditions. A total of 60 drivers (30 younger: 21-25 years, and older: 60-75 years old) are recruited for this driving simulator study, which is conducted as part of the EPSRC-funded HAROLD (HAzards, ROad Lighting and Driving) project. Response time and hit rate data for the DRT, the percentage of correct responses for the n-back task, as well as lateral and longitudinal vehicle metrics are collected, to understand how different lighting conditions affect stimulus detection, and how this is affected by engagement in a demanding cognitive distraction task. Results will be discussed in terms of the implications of such non-visually distracting tasks on driving performance, and road safety. The ultimate aim of the project is to understand how/if pedestrian detection at night is affected by driver engagement in cognitively demanding, non-visual, tasks, to contribute to state of the art on distraction and lighting research, together with policy and countermeasure development.

1. Introduction

1.1 Driver Distraction

Despite a continued and sustained effort to prevent the adverse effects of distraction while driving in recent years, distracted driving still appears to be a critical contributor to crash involvement (e.g. Lym & Chen, 2021; Olsson et al., 2020), perhaps due to the plethora of activities now possible on our mobile devices in the vehicle, as well as the general pressures of life, taking our minds off the main driving task. During the past 20 years, numerous studies have examined the effects of engagement in distracting vehicle-based activities on driving performance, such as how they divert our attention away from the driving task, and increase our brake response, and crash involvement (e.g. Li et al., 2019; Papantoniou et al., 2017).

Such distracting activities can be broadly categorised into those that require visual, visual-manual, auditory, and cognitive resources, or a combination of the above (NHTSA, 2010; Ranney et al., 2000). “Cognitive distraction”, which is associated with increased cognitive activity, includes thinking about something other than driving, taking attention and mind off the road (NHTSA, 2010). One lab-based task that has been used extensively to study the effect of cognitive distraction on driving performance, is the n-back task (Mehler et al., 2011; Stojmenova & Sodnik, 2018). Increased n-back difficulty is associated with increased cognitive load (Čegovnik et al., 2018), and a reliable measure for studying the effects of varying levels of cognitive load on driving outcomes (von Janczewski et al., 2021).

1.2 Night-Time Driving

The time of day is an important factor that might directly or indirectly affect driving outcomes through environmental factors such as visibility (Wood, 2020), and

exposure to different levels of risk (Åkerstedt et al., 2001). Driving at night is perceived to be riskier and more difficult compared to daytime driving, due to decreased visibility of the environment (Evans et al., 2020), as well as the likelihood of driving while sleepy (Chipman & Jin, 2009). In the UK, night-time driving is shown to be particularly problematic for young and middle-aged drivers, with a higher proportion of accidents with fatal injuries occurring at night, when compared to day-time driving (Regev et al., 2018).

1.3 Aim of the Present Study

The detection-response task is a standard measurement adopted by the International Organization for Standardization (ISO 17488:2016) to determine the attentional demands due to the cognitive load of a secondary task (ISO, 2016; Stojmenova & Sodnik, 2018). Changes in cognitive load can be assessed with DRT performance, in terms of both response time and hit rate (ISO, 2016). Drivers’ DRT performance is known to be affected by engagement in secondary tasks (Bowden et al., 2019), and influenced by driver- and environment-related factors (e.g. Engström et al., 2005; van Winsum, 2018). However, to the best of our knowledge, very little research is done on how the detection of objects in the driving scene is affected by different lighting conditions. In light of this research gap, the present study focuses on how young and older drivers’ DRT is affected by a cognitive distraction task, and whether different lighting conditions influence this performance.

2. Method

2.1 Participants

The data collection for this study is currently underway. A total of 60 drivers are signed up for participation, with the sample being equally distributed across two age

groups: young drivers (21-25 years old), and older drivers (60-75 years old). Gender is also balanced for each age group.

2.2 Materials and Tasks

Driving environment and lighting level: The study will be conducted in the University of Leeds Driving Simulator (UoLDS). The scenario contains a two-lane, contraflow, rural road, with a 60 mph speed limit, consisting of straight and curved road sections. The lighting of the driving simulator environment is presented at two levels (daylight and night-time).

The n-back task: An auditory version of the n-back task, first used Mehler et al., (2011), will be used to provide two levels of difficulty in cognitive distraction: 1-back (repeating the digit one before) and 2-back (repeating the digit two before) the last digit heard. Participants will be required to respond to an auditory stimulus, presented via the driving simulator speakers, and response is provided verbally, and recorded by the experimenter, and via a voice recorder. Each trial will include a set of randomly generated ten digits. The percentage of correct responses to the task will be used as an indication of n-back performance.

Detection-response task: The effect of the n-back task on cognitive load will be examined by using the visual DRT. Each trial will include the presentation of a red circle with a visual angle of about 1°, presented in the driving scene, at a horizontal angle of 11° to 23°, and a vertical angle of 2° to 4° above the horizon, on either the right or left side of the road used for the driving environment. Based on the ISO recommendations, these visual stimuli will be presented at a random rate of every three to five seconds. Participants will be asked to respond to the stimuli as quickly and accurately as possible by pressing a micro-switch button, which will be attached to the index finger of their dominant hand, against the steering wheel. Response time and hit rate will be calculated to evaluate performance (ISO, 2016).

2.3 Procedure

The study is approved by the ethics committee of the School of Business, Environment and Social Services, University of Leeds (AREA 21-108). After receiving informed consent and instructions, participants will first complete a practice drive, followed by practicing both the n-back and DRT, separately. They will then complete practice of driving with the n-back task, driving with the DRT, and driving with the n-back and DRT together. Following this practice drive, participants will complete two experimental drives, which will be exactly the same as the practice drive, and identical in terms of road geometry and presentation of the non-driving related tasks, but counterbalanced across participants, in terms of night- and day-time driving environment. Each of the non-driving related tasks are programmed to start in the straight sections of the road, and last around 30 seconds each. The total experiment duration, including familiarisation, briefing and subjective feedback takes approximately 60 minutes to complete, and participants will be compensated £20 for taking part in the study.

3. Results

Data collection is currently underway and results will be reported in the next version of this paper. Response time and the number of hits and misses to the DRT will be

calculated for the free (baseline) driving sessions with no n-back task, and compared to sections which require performance of the 1- and 2-back tasks. The effect of lighting conditions on detection of the stimuli will be studied and response from older and younger drivers will be compared, using mixed model ANOVAs.

4. Discussion

The findings of this study will be discussed, and the potential implications on road safety research and design practices will be outlined. The implications of these results on detection of pedestrians at night by distracted drivers will also be considered.

5. Acknowledgments

The HAROLD project (HAzards, ROad Lighting and Driving) was supported by the Engineering and Physical Sciences Research Council (EPSRC: EP/S003576/1). We would like to thank Anthony Horrobin for developing the driving simulator scenarios, and implementing the DRT and n-back tasks.

References

- Åkerstedt, T., Kecklund, G., & Hörte, L. G. (2001). Night driving, season, and the risk of highway accidents. *Sleep, 24*(4), 401–406. <https://doi.org/10.1093/sleep/24.4.401>
- Bowden, V. K., Loft, S., Wilson, M. D., Howard, J., & Visser, T. A. (2019). The long road home from distraction: Investigating the time-course of distraction recovery in driving. *Accident Analysis & Prevention, 124*, 23–32. <https://doi.org/10.1016/j.aap.2018.12.012>
- Čegovnik, T., Stojmenova, K., Jakus, G., & Sodnik, J. (2018). An analysis of the suitability of a low-cost eye tracker for assessing the cognitive load of drivers. *Applied Ergonomics, 68*, 1–11. <https://doi.org/10.1016/j.apergo.2017.10.011>
- Chipman, M., & Jin, Y. L. (2009). Drowsy drivers: The effect of light and circadian rhythm on crash occurrence. *Safety Science, 47*(10), 1364–1370. <https://doi.org/10.1016/j.ssci.2009.03.005>
- Engström, J., Johansson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*(2), 97–120. <https://doi.org/10.1016/j.trf.2005.04.012>
- Evans, T., Stuckey, R., & Macdonald, W. (2020). Young drivers' perceptions of risk and difficulty: Day versus night. *Accident Analysis & Prevention, 147*, 105753. <https://doi.org/10.1016/j.aap.2020.105753>
- International Organisation for Standardisation (ISO) (2016). ISO 17488:2016 – *Road vehicles - Transport information and control systems - Detection-response task (DRT) for assessing attentional effects of cognitive load in driving*. International Organization for Standardization. <https://www.iso.org/standard/59887.html>
- Li, X., Oviedo-Trespalacios, O., Rakotonirainy, A., & Yan, X. (2019). Collision risk management of cognitively distracted drivers in a car-following situation. *Transportation Research Part F: Traffic Psychology and Behaviour, 60*, 288–298. <https://doi.org/10.1016/j.trf.2018.10.011>

Lym, Y., & Chen, Z. (2021). Influence of built environment on the severity of vehicle crashes caused by distracted driving: A multi-state comparison. *Accident Analysis & Prevention, 150*, 105920. <https://doi.org/10.1016/j.aap.2020.105920>

Mehler, B., Reimer, B., & Dusek, J.A. (2011). *MIT AgeLab delayed digit recall task (n-back)*. MIT AgeLab White Paper Number 2011-3B. Massachusetts Institute of Technology, Cambridge, MA. http://web.mit.edu/reimer/www/pdfs/Mehler_et_al_n-back-white-paper_2011_B.pdf

NHTSA. (2010). Overview of the National Highway Traffic Safety Administration's Driver Distraction Program. https://one.nhtsa.gov/staticfiles/nti/distracted_driving/pdf/811299.pdf

Olsson, B., Pütz, H., Reitzug, F., & Humphreys, D. K. (2020). Evaluating the impact of penalising the use of mobile phones while driving on road traffic fatalities, serious injuries and mobile phone use: a systematic review. *Injury Prevention, 26*(4), 378–385. <http://dx.doi.org/10.1136/injuryprev-2019-043619>

Papantoniou, P., Papadimitriou, E., & Yannis, G. (2017). Review of driving performance parameters critical for distracted driving research. *Transportation Research Procedia, 25*, 1796–1805. <https://doi.org/10.1016/j.trpro.2017.05.148>

Ranney, T. A., Garrott, W. R., & Goodman, M. J. (2001). NHTSA driver distraction research: Past, present, and future (No. 2001-06-0177). SAE Technical Paper.

Regev, S., Rolison, J. J., & Moutari, S. (2018). Crash risk by driver age, gender, and time of day using a new exposure methodology. *Journal of Safety Research, 66*, 131–140. <https://doi.org/10.1016/j.jsr.2018.07.002>

Stojmenova, K., & Sodnik, J. (2018). Detection-response task—uses and limitations. *Sensors, 18*(2), 594. <https://doi.org/10.3390/s18020594>

van Winsum, W. (2018). The effects of cognitive and visual workload on peripheral detection in the detection response task. *Human Factors, 60*(6), 855–869. <https://doi.org/10.1177%2F0018720818776880>

von Janczewski, N., Wittmann, J., Engeln, A., Baumann, M., & Krauß, L. (2021). A meta-analysis of the n-back task while driving and its effects on cognitive workload. *Transportation Research Part F: Traffic Psychology and Behaviour, 76*, 269–285. <https://doi.org/10.1016/j.trf.2020.11.014>

Wood, J. M. (2020). Nighttime driving: visual, lighting and visibility challenges. *Ophthalmic and Physiological Optics, 40*(2), 187–201. <https://doi.org/10.1111/opo.12659>

Key requirements and a method for measuring in-vehicle user interfaces' distraction potential

Tuomo Kujala¹, Hilikka Grahn¹

¹ Cognitive science, Faculty of Information Technology, University of Jyväskylä, PO Box 35 FI-40014 University of Jyväskylä, Jyväskylä, Finland

*tuomo.kujala@jyu.fi

Numerous studies have shown the harmful effects of drivers' smartphone use on driver distraction and crash risk. However, there is less scientific research available on embedded in-car infotainment systems' distraction potential – even if drivers can nowadays use these systems to conduct highly similar tasks as with their smartphones. Unfortunately, the scientific community does not have an agreed-upon definition for driver inattention which has led to various operationalizations of inattention in distraction research. The lack of common definition and the different operationalizations hinder the comparability of distraction testing results. To guide the development of valid distraction testing procedures, this paper argues for eight key requirements for the operationalization of driver inattention: 1) Evaluation of inattention against attentive task performance, 2) Consideration of spare attentional capacity available in attentive driving, 3) Consideration of individual and situational differences in this spare attentional capacity, 4) Consideration of drivers' cognitive processing abilities, 5) Focusing evaluations on cognitive processes relevant for attentive driving, 6) Probabilistic evaluations in order to avoid hindsight bias, 7) Link to real-life crash risk – or real-life performance failure probability, and 8) Prioritization of possibility over probability. Example of a testing method fulfilling these requirements is illustrated for measuring in-vehicle user interfaces' distraction potential. The method aims at well-controlled, reliable, valid, and comparable testing of the distraction potential of infotainment systems. A practical aim of the testing is to encourage car manufactures to design safer user interfaces for their infotainment systems.

1. Introduction

A great number of studies have indicated the detrimental impacts of smartphone use on driving performance (e.g., Caird et al., 2014, 2018; Ferdinand & Menachemi, 2014; Guo et al., 2010; Horrey, 2018; Lipovac et al., 2017; Oviedo-Trespalacios et al., 2016; Papantoniou et al., 2017; Simmons et al., 2016, 2017). However, latest in-car infotainment systems are nowadays so advanced that they offer almost as wide range of applications as do smartphones. Unfortunately, scientific knowledge of the distraction potential of these original equipment manufacturer infotainment systems (OEM) is scarce.

There are varying authoritative recommendations and standards (e.g., ISO, 2017; 2016; NHTSA, 2013) regarding how to measure distraction potential of infotainment systems. All of these have faced criticism in the scientific community. These disagreements can lead to situations where it is difficult to interpret and compare different research outcomes (Lee et al., 2009; Pettitt et al., 2005; Regan et al., 2011). Often the measurements focus on the so-called visual or cognitive demands of the in-car tasks but it is unclear what is the relationship between these demands and their distraction potential (Grahn & Kujala, 2020). Especially, if there is no driving involved, how valid are these measurements for studying the in-car tasks' effects on driving?

Overall, when the distraction potential of in-car activities is studied, it is crucially important that we are actually measuring what we state to measure. Hence, we provide evidence-based arguments for eight key requirements for the operationalization of driver inattention to serve the development of valid and reliable testing measures. Finally,

we introduce an example of a testing method that could fulfill the requirements.

2. Key requirements for the measurement of in-vehicle user interfaces' distraction potential

1) *Inattention has to be evaluated against attentive task performance.* Driver inattention can be defined as insufficient, or no attention, to activities critical for safe driving (Regan et al., 2011). Driver distraction can then be defined as a specific form of inattention, which is caused by attention being diverted towards activities not related to driving, such as in-car tasks. The crucial point is that – by definition – inattention should be defined and measured against “activities critical for safe driving”. Hence, there should be a baseline of attentive driving to enable evaluation of how much this is distracted due to the in-car tasks (see, e.g., Grahn, 2021). This may sound obvious but, nevertheless, it has been proposed that distraction potential of in-car tasks could be assessed by, for instance, evaluating task durations under occlusion (e.g., NHTSA, 2013), without even simulating the driving task.

2) *Spare attentional capacity available in attentive driving has to be considered.* There is evidence showing that drivers can drive safely even if not allocating 100% attention continuously to the driving task (e.g., Kujala et al., 2021). In driving, there are safety-critical thresholds, such as lane boundaries and the rear bumper of the lead car. For attentive driving, the certainty that the car will stay within the associated safety margins – even if the driver is not paying attention at the time to the driving-relevant targets – is what matters. Therefore, spare attentional capacity and inattention should be judged in relation to these task-critical thresholds (Kujala, 2021).

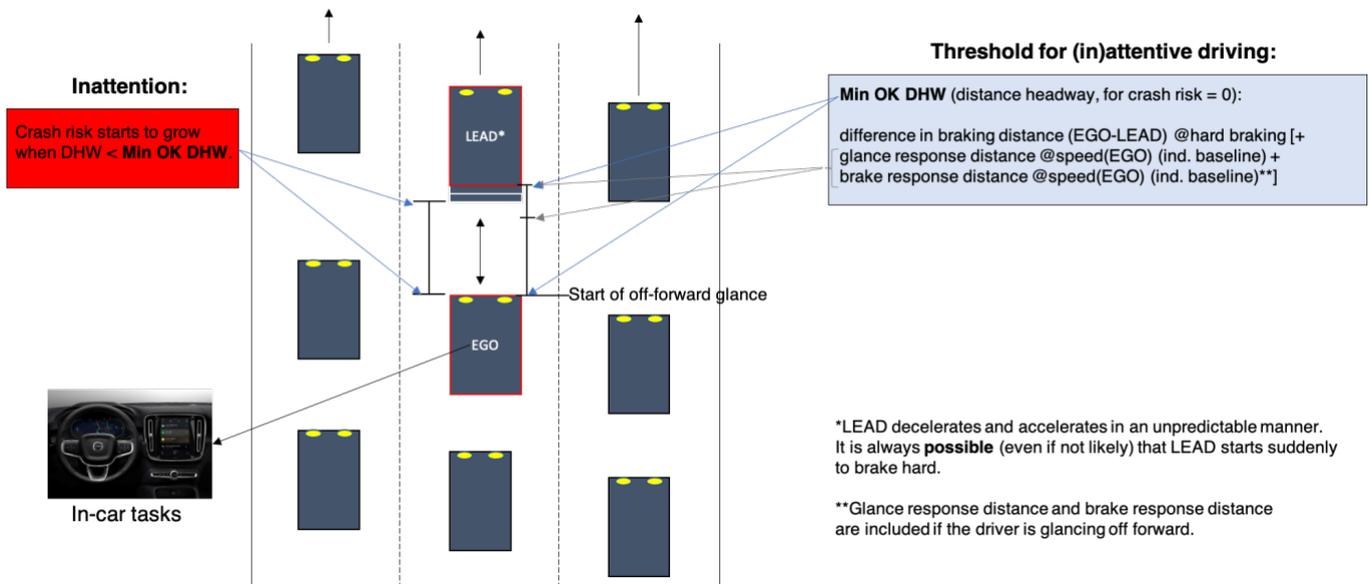


Fig. 1. A bird-view visualization of the proposed driving scenario and inattention measurement.

3) *Individual and situational differences in the spare attentional capacity have to be considered.* There are inter-individual (Broström et al., 2013, 2016; Chen & Donmez, 2016; Donmez et al., 2010; Grahn & Taipalus, 2021; Yang et al., 2021) and situational (e.g., Kujala et al., 2021; Large et al., 2015) differences in the spare attentional capacity in driving. This variability needs to be taken into account in testing to avoid, for instance, a situation where the characteristics of the test driver sample affect the test results more than the in-car user interfaces (Broström et al., 2016; Grahn & Taipalus, 2021; Ljung Aust et al., 2015). Furthermore, the threshold of inattention cannot be for each driver and situation a static, for instance, 2-second glance duration off forward (e.g., NHTSA, 2013). What matters for the distraction potential of in-car glances, is their timing in relation to safety-critical thresholds, and not their duration, *per se*.

4) *Drivers' cognitive processing abilities and limitations have to be considered.* If the driver can react in time to brake lights in front, as well as to sufficiently process all the other driving-related subtasks, it is not justified to claim that a driver is distracted when glancing upon an in-car display or a HUD. When measuring distraction, focal vision is often in a key position, but drivers can gather lots of information also peripherally (Ahlström et al., 2021; Svärd et al., 2021). Additionally, there are forms of cognitive distraction that do not display themselves via glance targets (e.g., internalized thoughts, Regan et al., 2011).

5) *Evaluation should focus on cognitive processes relevant for attentive driving.* Some cognitive processes are highly relevant to succeed in the driving task and some are not that relevant. Hence, measurements of distraction should focus on the in-car task's interference effects on those cognitive functions that are relevant in driving in real traffic, such as distance and speed estimations and hazard prediction. These should not focus on, for instance, off-forward glance durations or in-car task times (e.g., NHTSA, 2013) without considering how these relate to the cognitive processes that are actually relevant for real-world driving. For instance, according to a study by Nilsson et al. (2018) it is inappropriate to generalize from delayed response times in artificial

response tasks (e.g., detection response tasks, DRT) to more realistic safety-critical events in real traffic, such as looming of a lead car, for which responses are known to be triggered automatically.

6) *Evaluation has to be probabilistic in order to avoid hindsight bias.* Regan et al. (2011) and Kircher and Ahlström (2017) raise the problem of hindsight bias in defining inattention based on the outcome of a situation, such as, a crash. Instead, it should already be known before the outcome of an in-car activity if the driver is attentive or not. Therefore, we should be able to assess the possibility of a crash or performance failure in each situation – even if these would not realize (Kujala, 2021).

7) *There should be a link to real-life crash risk – or real-life performance failure probability.* This requirement is for ensuring the ecological relevance of the testing. Based on the distraction test results, we should be able to estimate what are the potential safety consequences or effects of the in-car activities on crash risk or performance failure probability in real traffic, even if the testing would be done in a driving simulator (Bärgman & Victor, 2020).

8) *Possibility should be more important than probability.* Skilled, attentive driving is about maintenance of appropriate safety margins. A critical component of this skill is the recognition of possible hazards ahead and acting upon them to prevent these, even if the potential hazards are not likely to happen (Grahn et al., 2020).

3. Method: An example on how to measure inattention by fulfilling the eight key requirements

In an example method (see Fig. 1), a test participant would drive the EGO vehicle surrounded by traffic and conduct in-car tasks, such as searching for music. The LEAD car in front decelerates and accelerates in an unpredictable manner. It is always possible – even if not likely – that the LEAD starts hard braking. The variable acceptable thresholds for the distance headway (DHW) are based on this possible hazardous scenario in all situations. The idea is that an attentive driver will always keep the rear-end crash risk at zero by adjusting one's speed and off-forward glancing behavior based on the variable situational factors, that is, the

distance and relative speed to the LEAD car. Here, the assessment would be focused on the impact(s) of the in-car tasks on maintenance of appropriate safety margins from the perspective of the possibility of a rear-end crash (i.e., is DHW smaller than minimum DHW to avoid crash). If the DHW to the LEAD grows too large for meaningful measurements, a car from an adjacent lane can change lane to become a new LEAD car.

In the method, two baseline drives of attentive driving in a similar scenario are required: a) One for measuring driver's minimum and comfortable preferences for DHW while keeping eyes on the LEAD car (Taieb-Maimon & Shinar, 2001) and b) another for measuring driver's maximum glance and brake response times (Svärd et al., 2021) when eyes are off the LEAD car when it suddenly brakes hard. For comparability, the test participant samples between tests should be balanced so that there would be a similar representation of drivers who prefer shorter and drivers who prefer longer DHWs. Alternatively, other means should be utilized to control for these inter-individual differences in test results.

Currently, we are building a testing environment, a drive-in lab, in which we can drive in a commercial vehicle, connect it into our driving simulation and test the OEM infotainment system of the car in a driving scenario with the metrics described in Figure 1. We are first targeting a benchmarking study with popular 2022 car models. We expect that the final results are ready before the full paper's submission deadline.

4. Conclusions

Here, we have presented eight key requirements and a testing method for measuring in-vehicle user interfaces' distraction potential that could fulfill these requirements. The requirements aim at reliable, valid and comparable testing on the distraction potential of infotainment systems. These can be also utilized in evaluating existing guidelines and proposed methods for distraction testing. A practical aim of this methodological development and testing is to urge car manufactures to design safer user interfaces for their infotainment systems. The requirements are open for scientific debate, for instance, the list might not be comprehensive.

Acknowledgements

The research was partly funded by Academy of Finland (Appropriate Uncertainty in Manual and Automated Driving/343259).

References

- Ahlström, C., Kircher, K., Nyström, M., & Wolfe, B. (2021). Eye tracking in driver attention research: How gaze data interpretations influence what we learn. *Frontiers in Neuroergonomics*, *2*, 778043. <https://doi.org/10.3389/fnrgo.2021.778043>
- Bärgman, J., & Victor, T. (2020). Holistic assessment of driver assistance systems: How can systems be assessed with respect to how they impact glance behaviour and collision avoidance? *IET Intelligent Transport Systems*, *14*(9), 1058–1067. <https://doi.org/10.1049/iet-its.2018.5550>
- Broström, R., Aust, M. L., Wahlberg, L., & Källgren, L. (2013). What drives off-road glance durations during multitasking – capacity, practice or strategy? *Proceedings on the 3rd International Conference on Driver Distraction and Inattention*.
- Broström, R., Bengtsson, P., & Aust, M. L. (2016). Individual glance strategies and their effect on the NHTSA visual manual distraction test. *Transportation Research Part F: Traffic Psychology and Behaviour*, *36*, 83–91. <https://doi.org/10.1016/j.trf.2015.10.017>
- Caird, J. K., Johnston, K. A., Willness, C. R., Asbridge, M., & Steel, P. (2014). A meta-analysis of the effects of texting on driving. *Accident Analysis and Prevention*, *71*, 311–318. <https://doi.org/10.1016/j.aap.2014.06.005>
- Caird, J. K., Simmons, S. M., Wiley, K., Johnston, K. A., & Horrey, W. J. (2018). Does talking on a cell phone, with a passenger, or dialing affect driving performance? An updated systematic review and meta-analysis of experimental studies. *Human Factors*, *60*(1), 101–133. <https://doi.org/10.1177/0018720817748145>
- Chen, H.-Y. W., & Donmez, B. (2016). What drives technology-based distractions? A structural equation model on social-psychological factors of technology-based driver distraction engagement. *Accident Analysis & Prevention*, *91*, 166–174. <https://doi.org/10.1016/j.aap.2015.08.015>
- Donmez, B., Boyle, L. N., & Lee, J. D. (2010). Differences in off-road glances: Effects on young drivers' performance. *Journal of Transportation Engineering*, *136*(5), 403–409. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000068](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000068)
- Ferdinand, A. O., & Menachemi, N. (2014). Associations between driving performance and engaging in secondary tasks: A systematic review. *American Journal of Public Health*, *104*(3). <https://doi.org/10.2105/AJPH.2013.301750>
- Grahn, H. (2021). *On the measurement of visual distraction potential of in-car activities*. JYU Dissertations 430. University of Jyväskylä.
- Grahn, H., & Kujala, T. (2020). Impacts of touch screen size, user interface design, and subtask boundaries on in-car task's visual demand and driver distraction. *International Journal of Human Computer Studies*, *142*. <https://doi.org/10.1016/j.ijhcs.2020.102467>
- Grahn, H., Kujala, T., Silvennoinen, J., Leppänen, A., & Saariluoma, P. (2020). Expert drivers' prospective thinking-aloud to enhance automated driving technologies – Investigating uncertainty and anticipation in traffic. *Accident Analysis and Prevention*, *146*, 105717. <https://doi.org/10.1016/j.aap.2020.105717>
- Grahn, H., & Taipalus, T. (2021). Refining distraction potential testing guidelines by considering differences in glancing behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, *79*, 23–34. <https://doi.org/10.1016/j.trf.2021.03.009>
- Guo, F., Klauer, S. G., Hankey, J. M., & Dingus, T. A. (2010). Near crashes as crash surrogate for naturalistic driving studies. *Transportation Research Record: Journal of the Transportation Research Board*,

- 2147(1), 66–74. <https://doi.org/10.3141/2147-09>
- International Organisation for Standardization. (2017). *Road vehicles — Ergonomic aspects of transport information and control systems — Occlusion method to assess visual demand due to the use of in-vehicle systems (ISO Standard No. 16673)*.
- International Organisation for Standardization. (2016). *Road vehicles — Transport information and control systems — Detection-response task (DRT) for assessing attentional effects of cognitive load in driving (ISO Standard No. 17488)*.
- Kircher, K., & Ahlström, C. (2017). Minimum required attention: A human-centered approach to driver inattention. *Human Factors*, 59(3), 471–484. <https://doi.org/10.1177/0018720816672756>
- Kujala, T. (2021). Quantifying attentional demand of a lane-keeping task as the minimum required information in predictive processing. *7th International Conference on Driver Distraction and Inattention: Book of Abstracts*, 54–57. https://ddi2020.sciencesconf.org/data/pages/Proceedings_DDI_211021.pdf
- Kujala, T., Kircher, K., & Ahlström, C. (2021). A review of occlusion as a tool to assess attentional demand in driving. *Human Factors*. <https://doi.org/10.1177/00187208211010953>
- Large, D. R., van Loon, E., Burnett, G., & Pournami, S. (2015). Applying NHTSA task acceptance criteria to different simulated driving scenarios. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 117–124. <https://doi.org/10.1145/2799250.2799254>
- Lee, J. D., Young, K. L., & Regan, M. A. (2009). Defining driver distraction. In M. A. Regan, J. D. Lee, & K. L. Young (Eds.), *Driver Distraction – Theory, Effects, and Mitigation* (pp. 31–40). CRC Press Taylor & Francis Group.
- Lipovac, K., Đerić, M., Tešić, M., Andrić, Z., & Marić, B. (2017). Mobile phone use while driving - Literary review. *Transportation Research Part F: Traffic Psychology and Behaviour*, 47, 132–142. <https://doi.org/10.1016/j.trf.2017.04.015>
- Ljung Aust, M., Rydstöm, A., Broström, R., & Victor, T. (2015). Better ways to calculate pass/fail criteria for the eye glance measurement using driving simulator test. *4th International Driver Distraction and Inattention Conference*.
- National Highway Traffic Safety Administration. (2013). Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices. In *Docket No. NHTSA-2010-0053*. <https://doi.org/NHTSA-2010-0053>
- Nilsson, E. J., Aust, M. L., Engström, J., Svanberg, B., & Lindén, P. (2018). Effects of cognitive load on response time in an unexpected lead vehicle braking scenario and the detection response task (DRT). *Transportation Research Part F: Traffic Psychology and Behaviour*, 59, 463–474.
- Oviedo-Trespalacios, O., Haque, M. M., King, M., & Washington, S. (2016). Understanding the impacts of mobile phone distraction on driving performance: A systematic review. *Transportation Research Part C: Emerging Technologies*, 72, 360–380. <https://doi.org/10.1016/J.TRC.2016.10.006>
- Papantoniou, P., Papadimitriou, E., & Yannis, G. (2017). Review of driving performance parameters critical for distracted driving research. *Transportation Research Procedia*, 25, 1796–1805. <https://doi.org/10.1016/j.trpro.2017.05.148>
- Pettitt, M., Burnett, G., & Stevens, A. (2005). Defining driver distraction. *World Congress on Intelligent Transportation Systems*.
- Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis and Prevention*, 43(5), 1771–1781. <https://doi.org/10.1016/j.aap.2011.04.008>
- Simmons, S. M., Caird, J. K., & Steel, P. (2017). A meta-analysis of in-vehicle and nomadic voice-recognition system interaction and driving performance. *Accident Analysis and Prevention*, 106, 31–43. <https://doi.org/10.1016/j.aap.2017.05.013>
- Simmons, S. M., Hicks, A., & Caird, J. K. (2016). Safety-critical event risk associated with cell phone tasks as measured in naturalistic driving studies: A systematic review and meta-analysis. *Accident Analysis and Prevention*, 87, 161–169. <https://doi.org/10.1016/j.aap.2015.11.015>
- Svärd, M., Bårgman, J., & Victor, T. (2021). Detection and response to critical lead vehicle deceleration events with peripheral vision: Glance response times are independent of visual eccentricity. *Accident Analysis & Prevention*, 150, 105853. <https://doi.org/10.1016/J.AAP.2020.105853>
- Taieb-Maimon, M., & Shinar, D. (2001). Minimum and comfortable driving headways: Reality versus perception. *Human Factors*, 43(1), 159–172.
- Yang, S., Kuo, J., & Lenné, M. G. (2021). Individual differences in glance patterns under distraction in level 2 automated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. <https://doi.org/https://doi.org/10.1177/1071181320641477>

Quantification of driver's side-glance frequency and duration in straight highway driving

Pierluigi Olleja¹, Jonas Bärgrman¹, Gustav Markkula²

¹ Division of Vehicle Safety, Chalmers University of Technology, Hörselgången 4, Gothenburg, Sweden

² Institute for Transport Studies, University of Leeds, University Road 34-40, Leeds, UK

* pierluigi.olleja@chalmers.se

Abstract: Reference driver models as safety targets for the virtual assessment of autonomous vehicles (AVs) are currently being developed. The pilot study presented in this abstract aims to quantify drivers' side-glance behaviors in terms of frequency and duration, when driving on a highway without changing lanes. Before the conference presentation, the results will be used in virtual simulations to assess whether including these behaviors in a reference driver model would have any practical impact on the AV's safety assessment for cut-ins and side-swipe crash scenarios.

1. Introduction

Driving context has been shown to influence drivers' glance behavior (Morando et al., 2019). Additionally, off-path glance duration has been shown to have an impact on the probability of a crash or near-crash (Tian et al., 2013; Victor et al., 2015). Many studies have focused on the duration of on-path and off-path glances in rear-end conflict scenarios (Klauer et al., 2014; Victor et al., 2015). Others have studied the use of mirrors prior to lane changes (Pech et al., 2014), and right turns (Jansen et al., 2017). Only rarely has research focused on the use and safety impact of side glances to proactively check potential threats moving laterally (e.g., cut-ins and side-swipes). Studies about the frequency, duration, and timing (as a function of context) of highway drivers' side glances (towards the side mirror and side window) are not currently available.

As Automated Vehicles (AVs) continue to be developed, there is a need to assess their safety impact. In the last few years, reference driver models are being increasingly used as safety targets for AVs (Webb et al., 2020). It is unclear whether including a model of side glances in reference driver models for cut-ins and side-swipes would impact the virtual safety assessment outcome in any practical way. The significance of including driver side-glance behaviour in such models can be assessed in three steps: 1) quantification of driver side-glance behaviors, 2) a model of driver avoidance manoeuvres when a threat is identified, and 3) virtual simulations to assess the safety impact of the combination of 1) and 2) on crash avoidance and mitigation. If the results show a substantial impact on crash avoidance and/or mitigation, a side-glance behavior model should be included in reference driver models assessing AVs for cut-ins and side-swipes.

1.1 Aims and objectives

The aim of this (pilot) work is to perform Step 1 above: quantifying, as a function of context, the frequency and duration of drivers' side mirror and side window glances in no-lane-change (undisturbed) highway driving. The context is considered in order to better understand the overall variability of side-glance behaviors for future modeling. Before the conference, we will create a simple avoidance response model (Step 2: likely based on SHRP2 naturalistic

driving data). We will then perform virtual simulations (Step 3) on reconstructed cut-in and side-swipe crashes with and without the side-glance behavior model. Comparing the results will allow us to assess whether the model impacts the crash rate or impact speeds. We may also collect additional side-glance behavior data for use in the simulations.

2. Method

2.1 Data description

The data used in this study were collected during the L3Pilot project. The full dataset includes trips with manually driven vehicles (baseline) and trips in which an automated driving function was deactivated. In this study, only baseline trips with non-professional drivers were analysed. The vehicles were equipped with five cameras recording the driver, the forward and rear views, and the cockpit from various angles. Surrounding objects were automatically detected and used to determine the traffic density. Specific driving scenarios (e.g., car following, lane change, and cut-in) were identified automatically. In this study, only the free-driving and car-following scenarios (without lane changes) were analyzed.

2.2 Definition of low and high traffic density

Data from 20 drivers were used in this pilot study. The main author manually annotated the driver glances in two 30-s long events of free driving: one each of low and high(er) traffic density. In low-density traffic events, the driver's vehicle was near, at most, one vehicle (in the same lane or an adjacent one). High-density traffic events were defined as having four or more vehicles nearby.

Event selection was accomplished by: a) using an algorithmically defined "TrafficDensity" metric, defined as "vehicle / km / lane", to define the lowest and highest traffic-density events for each driver and b) performing a visual check to exclude events with misclassifications.

2.3 Annotation of glances

Glances were primarily annotated using the videos recorded by the camera on the driver-side A pillar (facing the driver). Additional videos from other camera angles in the cockpit were used when glancing was unclear in the primary video. The following types of glances were annotated (along

with their duration): left-side mirror and window, right-side mirror and window, and rear-view mirror. In this pilot, the annotator included only focused glances toward the mirrors and windows (excluding side glances clearly due to distractions); the main study will handle glance selection differently.

3. Results

Fig. 1 the ratio between the side glances in high and low density traffic for each driver. Visually there seems to be a trend towards side-glances being more frequent in high traffic density than in low. However, with the small sample, neither a Wilcoxon rank sum test (of the difference between counts), nor a Poisson linear mixed effect model (predicting the number of glances with driver as random effect) showed significant differences in these pilot data.

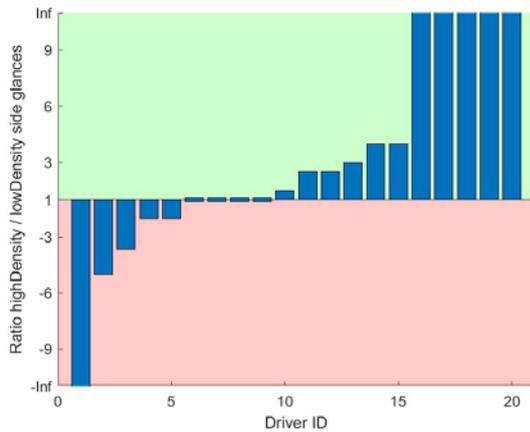


Fig. 1. Ratio between side glances in low and high traffic density. The green area represents a higher number of side glances in high traffic density; the red area, a higher number of side glances in low traffic density. For visualization purposes, ratios < 1 (when there were more side glances in low than high density traffic) were inverted, and a minus sign was added. Additionally, cases which presented zero in one of the two driving conditions were automatically assigned ± 10 in the bar plot.

Fig. 2 compares the glance durations for both traffic densities. The average duration of side glances was 0.79 s (SD ± 0.38 s) in low traffic density and 0.62 s (SD ± 0.29 s) in high traffic density. Additionally, the side glance frequency was computed for both traffic densities. The average frequency was 3.3 glances/minute in low traffic density and 4.1 glances/minute in high traffic density. The overall average was 3.7 glances/minute.

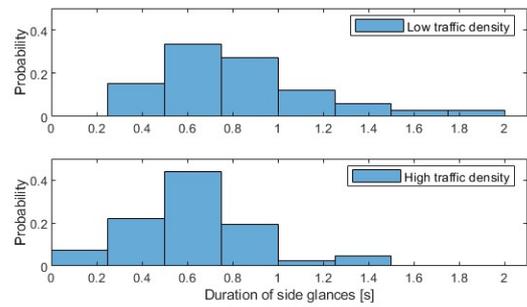


Fig. 2. The side-glance duration distributions for low and high traffic density data for the 20 drivers, respectively.

4. Discussion

4.1 Difference between low and high density

Our research shows a trend towards a larger number of driver side glances in higher density traffic. This trend could be a consequence of the driver’s need to know the positions and trajectories of the surrounding vehicles—information the driver uses to stay safe. It makes sense that an attentive driver would look more often at the mirrors when in higher density traffic, as the risk of conflicts may be higher.

Some drivers did not show a noticeable difference in behavior between low and high traffic density situations and some actually performed more side glances in the low traffic density scenario. These behaviors may be due to higher levels of driver confidence or to the peculiar behavior of a nearby vehicle that required more attention than “normal”.

4.2 Frequency and duration relevance for reference drivers

Reference driver models for virtual AV safety assessment lack a side glance model. If reference driver models for cut-ins and side-swipes were to include one, it might include a probabilistic description of how often and for how long an attentive, experienced, and defensive driver would check the side mirrors and side windows. The relevance of such a model is, however, unclear from a safety assessment perspective. That is, if the inclusion of a side-check (attention) model is unlikely to improve crash avoidance in the targeted scenarios, then keeping the reference driver models simple (excluding side-glance behavior) may be best. This study provided a preliminary estimate of the frequency and duration of driver side glances. The main study will run virtual simulations (likely with more data: a few more drivers and several more 30-s segments per driver) as well as a simple brake/acceleration response model based on SHRP2 data, providing a more robust assessment of the safety impact of the glance behavior model for side-swipe and cut-in scenarios.

These data can be used to simulate an attentive driver (using the upper end of the frequency/duration distributions) to use as a reference driver safety target for AVs. Similarly, the data could serve to simulate an inattentive driver (using the lower end of the frequency/duration distributions) to assess how that model’s performance compares with that of

the AV. Another possible application is creating virtual-simulation-based impact assessments to estimate the reduction of AV crashes/injuries compared to manual driving (using the full frequency/duration distributions).

5. Conclusions

The overall frequency of side glances was 3.7 glances/minute, with an average duration of 0.70 s (SD \pm 0.34 s). At this point, it is unclear if this is often (and long) enough to have an effect on reference driver models for the assessment of AVs' cut-in and side-swipe crash avoidance functionalities. However, we hope that the main study will answer this question.

6. Acknowledgments

We want to thank the the Swedish FFI funding program for funding the QUADRIS project, in which this research has been performed. r.

References

- Jansen, R. J., Lotan, T., Winkelbauer, M., Bärgrman, J., Kovaceva, J., Donabauer, M., Pommer, A., Musicant, O., Harel, A., Wesseling, S., Christoph, M., & van Nes, N. (2017). Interactions with vulnerable road users. UDRIVE Deliverable 44.1. https://doi.org/10.26323/UDRIVE_D44.1
- Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted Driving and Risk of Road Crashes among Novice and Experienced Drivers. *New England Journal of Medicine*, 370(1), 54–59. <https://doi.org/10.1056/NEJMs1204142>
- Morando, A., Victor, T., & Dozza, M. (2019). A Reference Model for Driver Attention in Automation: Glance Behavior Changes During Lateral and Longitudinal Assistance. *IEEE Transactions on Intelligent Transportation Systems*, 20(8), 2999–3009. <https://doi.org/10.1109/TITS.2018.2870909>
- Pech, T., Lindner, P., & Wanielik, G. (2014). Head tracking based glance area estimation for driver behaviour modelling during lane change execution. 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 655–660. <https://doi.org/10.1109/ITSC.2014.6957764>
- Tian, R., Li, L., Chen, M., Chen, Y., & Witt, G. J. (2013). Studying the Effects of Driver Distraction and Traffic Density on the Probability of Crash and Near-Crash Events in Naturalistic Driving Environment. *IEEE Transactions on Intelligent Transportation Systems*, 14(3), 1547–1555. <https://doi.org/10.1109/TITS.2013.2261988>
- Victor, T., Dozza, M., Bärgrman, J., Boda, C. N., Engström, J., Flannagan, C., Lee, J. D., & Markkula, G. (2015). Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk. In SHRP 2 Report S2-S08A-RW-1. <https://doi.org/doi:10.17226/22297>
- Webb, N., Smith, D., Ludwick, C., Victor, T., Hommes, Q., Favaro, F., Ivanov, G., & Daniel, T. (2020). Waymo's Safety Methodologies and Safety Readiness Determinations. <http://arxiv.org/abs/2011.00054>

Validation of an eyes-off-road crash causation model for virtual safety assessment

Jonas Bårgman^{1*}, Malin Svård^{1,2}, Simon Lundell², Ahmed Shams El Din²

¹ Chalmers University of Technology, Crash Analysis and Prevention Unit in the Division of Vehicle Safety, 412 95, Gothenburg, Sweden

² Volvo Car Corporation, Volvo Cars Safety Centre, 418 78 Gothenburg, Sweden

*jonas.bargman@chalmers.se

Abstract: Models of crash causation have been used in virtual simulations for some time, but such models are seldom thoroughly validated. This study aims to validate a virtual simulation-based rear-end crash causation model that uses two distributions of eyes-off-road glances (one from normal everyday driving and one from crashes) and a driver response model for rear-end crashes. Results show that the crash causation model does a reasonably good job of replicating the distribution of the impact speeds of the original crashes.

1. Introduction

Driver distraction and inattention (DDI) are common crash causation mechanisms (Klauer et al., 2014; NHTSA, 2018). Specifically, glances off-road have been shown to substantially increase crash and injury risks in manual driving (Victor et al., 2015). It is therefore natural to consider eyes-off-road glances (EOFF) as a crash causation mechanism, be it for the virtual safety assessment of advanced driver assistance systems (ADAS; Bårgman et al., 2017), of higher levels of automation (Bjorvatn et al., 2021), or of driver (glance) behavior changes—either associated with the introduction of new human-machine interfaces (HMI; Bårgman et al., 2015; Lee et al., 2018), or in combination with ADAS (Bårgman & Victor, 2020). These assessments typically implement EOFF-induced crash causation as mathematical models. However, such models have rarely been validated against real crash data.

1.1 Aims and objectives

The aim of this study is to validate a virtual simulation-based crash causation model that uses EOFF glance distributions and a driver response model for rear-end crashes. Specifically, two different EOFF glance behaviors were virtually applied to reconstructed crash kinematics. The impact speed distributions of the original (reconstructed) crashes were compared with those from the simulations.

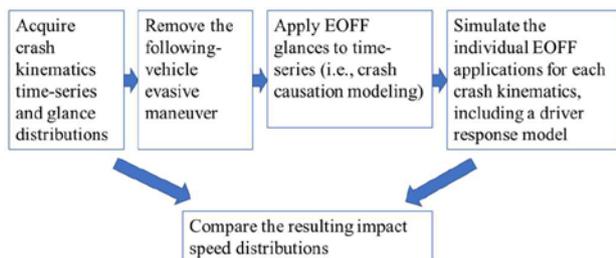


Figure 1: The workflow used to validate the EOFF crash causation and driver response model against reconstructed rear-end crash kinematics.

2. Method

2.1 Overview

This study’s workflow is illustrated in Figure 1.

2.2 Data and evasive maneuver removal

The study used the time-series positions of both the lead vehicle (LV) and the following vehicle (FV) from 44 reconstructed rear-end crashes in Volvo Cars’ crash database.

As crash reconstructions include several assumptions, the pre-crash kinematics are just estimates of the actual kinematics. The impact of these assumptions on the results was reduced by creating synthetic variations (e.g., in speed, deceleration, and timing) of the reconstructed crashes. Using the same variations as in the L3Pilot project (Bjorvatn et al., 2021, p. A24), we created a total of eleven variants of each reconstructed crash (i.e., 484 sets of rear-end crash kinematics).

The FV driver’s evasive maneuver must be removed from each reconstruction so that the driver models in the simulations can replace it. The maneuver is removed by identifying its start and assuming the driver would have continued at the speed in the timestep just before it.

2.3 Eyes-off-road glances and their application

The crash causation model to be validated is based on the application of the FV driver’s EOFF glances to the crash kinematics. Here “application” refers to the placement of individual off-road glances of specific durations virtually in the time-series (Figure 2). When the driver looks back towards the road, the response model takes over control.

This study used two different eyes-off-road (EOFF) distributions (Figure 3): one baseline EOFF and one crash EOFF (see Victor et al., 2015, for a description of the glance extraction).

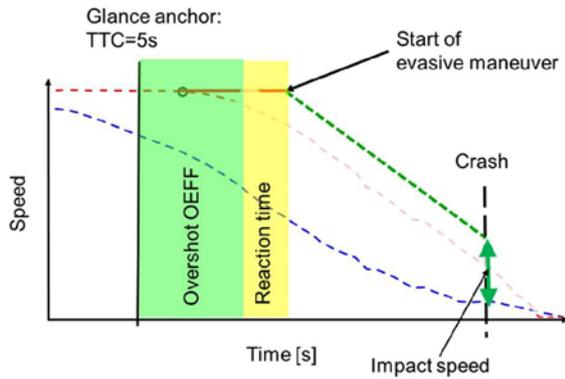


Figure 2: Illustration of the placement of an off-road (overshot) glance at a glance anchor on the kinematics of the LV and FV in a rear-end conflict. Dashed red line: the original FV speed; solid red line: the new LV speed after removal of the evasive maneuver; blue dashed line: the LV speed used; and green dashed line: applied full braking.

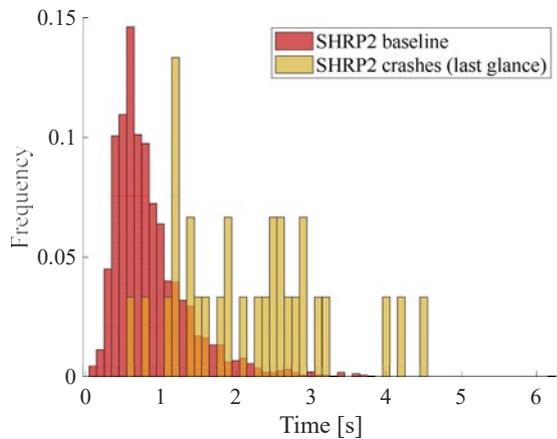


Figure 3: EOFF distributions.

As the kinematics are time series and the simulations play out in time, a model of how to place EOFF glances in the time series is needed. Following Bärgrman et al., (2015), in this study we assume that drivers do not start looking away from the roadway when the (optically defined) time to collision (TTC) is less than 5 s (based on Markkula et al., 2016), and that the probability of a lead vehicle braking is the same at all times.

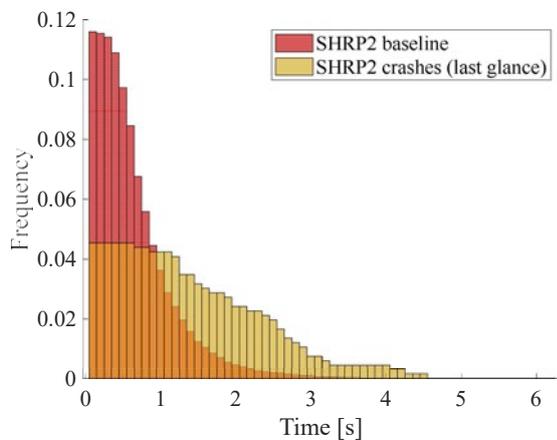


Figure 4: EOFF overshoot distributions.

An overshoot distribution with these assumptions can be applied to the EOFF distribution, generating an overshoot EOFF glance that can be placed at the last point in time with $TTC > 5$ s (Figures 2 and 4; see Bärgrman et al., 2017 for details).

2.4 Driver response model

A driver response model is needed to simulate the FV driver's response to the unfolding conflict (e.g., LV braking). This study uses a simple response model: when the driver glances back at the road, a reaction time of 0.5 s is added (based on Markkula et al., 2016) before (full) braking is applied.

2.5 Sampling and simulations

The distribution of impact speeds associated with each specific EOFF glance distribution was created by "applying" each EOFF distribution to each of the (484) crashes, so that every crash was simulated with each "bin" in the glance distribution.

2.6 Analysis

The impact speed distribution of the original 484 crashes was compared with that of all crashes in the simulations. This study used inverse-probability weighting on the EOFF overshoot distribution bins to minimize the number of simulations.

3. Results

Figure 5 shows the impact speed distributions of the simulated crashes generated by applying the crash causation model using glances from the SHRP2 EOFF distributions for baseline (upper panel) and for crashes (lower panel).

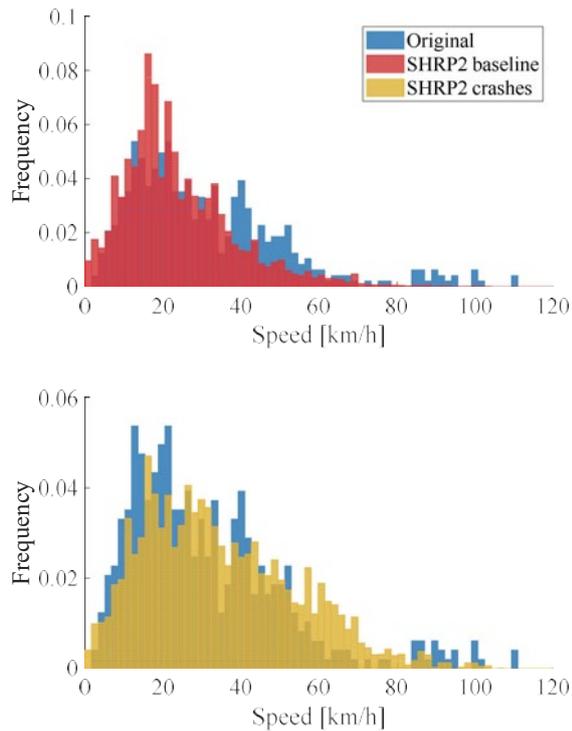


Figure 5: Impact speed distributions. *Upper:* Original impact speeds and impact speeds resulting from crash generation using the SHRP2 baseline EOFF distribution. *Lower:* Original impact speeds and impact speeds resulting from crash generation using the SHRP2 crashes EOFF distribution.

4. Discussion

4.1 Validation of the crash causation method

Results show relatively good similarity between the impact speeds of the original crashes and those of the crashes generated with the crash causation model in combination with the driver response model. The crash glances created a more similar impact speed distribution than the baseline glances but overestimated the proportion of crashes in the 60–80 km/h range. In contrast, the baseline EOFF distribution underestimated the proportion of crashes with an impact speed > 40 km/h. It is not surprising that crash glances result in higher impact speeds than baseline glances, since the crash EOFF distribution contains a higher proportion of long glance durations. However, both distributions underestimated the proportion of high-speed crashes (5 % occurred at 90–120 km/h). This may indicate that driver distraction or inattention (correlated with off-road glances) were not the underlying crash causation mechanism in these cases—it may instead have been drowsiness (possibly simulated as very long off-road glances/closed eyes).

4.2 Applications of simulations of crash causation mechanisms

There are several different uses for validated crash causation models; for example, investigating the crash and injury risks associated with the introduction of ADAS

through counterfactual simulations (Bärgman et al., 2015). Another use, which is receiving more and more attention, is in traffic simulations to “create” crashes (Bjorvatn et al., 2021). It is particularly important that any crash causation model used in traffic simulation is validated for each individual scenario, because crash generation is based not only on the stochastics of glance behavior, but also on many other factors that together might induce a crash (in contrast to typical counterfactual simulations). A third, albeit little studied or used, use of EOFF crash causation models is to assess glance behavior changes resulting from the introduction of some new HMI. The use of these models makes it possible to calculate a (continuous) safety metric for a specific HMI implementation (or, actually, the EOFF distribution associated with it). The metric is directly coupled to real-world crash causation through validation of the crash causation model (Bärgman et al., 2015).

4.3 Crash causation models

In this study, an EOFF-based crash causation model for rear-end crashes was validated. However, distraction/inattention leading to longer/more frequent off-road glances is only one crash causation mechanism among many (NHTSA, 2018). Driving while drowsy/sleepy and, for rear-end crashes, leaving too short a time gap to the lead vehicle are other common mechanisms. If virtual simulation-based safety benefit assessment is to gain acceptance, more crash causation models need to be developed, **and** validated.

4.4 Limitations

We used a relatively simple driver response model with a reaction threshold on looming and assumed maximum braking. Today there are better response models available (e.g., Svärd et al., 2017), but in this paper we show that even this simple model is a fairly good representation of driver responses to critical events.

5. Conclusions

Through virtual simulations, we have demonstrated that a mathematical crash causation model based on drivers’ off-road glances can generate an impact speed distribution similar to that of an in-depth crash database (rear-end crashes). This finding brings us a step closer to using such models for the virtual assessment of both ADAS and higher levels of automation, as well as for the virtual safety assessment of HMIs.

6. Acknowledgments

We would like to thank VINNOVA and Energimyndigheten for the partial funding we received for this work through the QUADRIS project (through FFI). We also want to thank Volvo Car Corporation and the Chalmers Area of Advance for their partial financing of this work. Fredrik Sangberg should also receive a special thank you for preparing the data. We further want to thank VTTI, the Transportation Research Board and the National Academies for making the SHRP2 data available. The findings and conclusions of this paper are those of the authors and do not necessarily represent the views of said organizations. Finally, we would like to thank Kristina Mayberry for her language review.

Bibliography

- Bärgman, J., Boda, C.-N., & Dozza, M. (2017). Counterfactual simulations applied to SHRP2 crashes: The effect of driver behavior models on safety benefit estimations of intelligent safety systems. *Accident Analysis & Prevention, 102*, 165–180. <https://doi.org/https://doi.org/10.1016/j.aap.2017.03.003>
- Bärgman, J., Lisovskaja, V., Victor, T., Flannagan, C., & Dozza, M. (2015). How does glance behavior influence crash and injury risk? A ‘what-if’ counterfactual simulation using crashes and near-crashes from SHRP2. *Transportation Research Part F: Traffic Psychology and Behaviour, 35*, 152–169. <https://doi.org/10.1016/j.trf.2015.10.011>
- Bärgman, J., & Victor, T. (2020). Holistic assessment of driver assistance systems: how can systems be assessed with respect to how they impact glance behaviour and collision avoidance? In *IET Intelligent Transport Systems* (Vol. 14, Issue 9, pp. 1058–1067). Institution of Engineering and Technology. <https://digital-library.theiet.org/content/journals/10.1049/iet-its.2018.5550>
- Bjorvatn, A., Page, Y., Fahrenkrog, F., Weber, H., Aittoniemi, E., Heum, P., Lehtonen, E., Silla, A., Bärgman, J., Borrack, M., Innamaa, S., Itkonen, T., Malin, F., Pedersen, K., Schuldes, Michael Sintonen, H., Streubel, T., Hagleitner, W., Hermitte, T., Hiller, J., & Torrao, G. (2021). *L3Pilot Deliverable D7.4 - Impact Evaluation Results*. https://l3pilot.eu/fileadmin/user_upload/Downloads/Deliverables/Update_14102021/L3Pilot-SP7-D7.4-Impact_Evaluation_Results-v1.0-for_website.pdf
- Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. *New England Journal of Medicine, 370*(1), 54–59. <https://doi.org/10.1056/NEJMsa1204142>
- Lee, J. Y., Lee, J. D., Bärgman, J., Lee, J., & Reimer, B. (2018). How safe is tuning a radio?: using the radio tuning task as a benchmark for distracted driving. *Accident Analysis & Prevention, 110*(Supplement C), 29–37. <https://doi.org/https://doi.org/10.1016/j.aap.2017.10.009>
- Markkula, G., Engström, J., Lodin, J., Bärgman, J., & Victor, T. (2016). A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies. *Accident Analysis & Prevention, 95, Part A*, 209–226. <https://doi.org/10.1016/j.aap.2016.07.007>
- NHTSA. (2018). *Traffic safety facts: Critical reasons for crashes investigated in the national motor vehicle crash causation survey*.
- Svärd, M., Markkula, G., Engström, J., Granum, F., & Bärgman, J. (2017). A quantitative driver model of pre-crash brake onset and control. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*(1), 339–343. <https://doi.org/10.1177/1541931213601565>
- Victor, T., Dozza, M., Bärgman, J., Boda, C.-N., Engström, J., Flannagan, C., Lee, J. D., Markkula, G., & Sciences, U. S. N. A. of. (2015). Analysis of naturalistic driving study data: SAFER glances, driver inattention, and crash risk. In *SHRP 2 safety project SO8* (Issue SHRP 2 SO8A). Transportation Research Board of the National Academy of Sciences. <https://doi.org/ISBN:978-0-309-27423-4>

Strategical use of peripheral vision in driving

Christer Ahlström^{1,2*}, Katja Kircher¹, Christian Vater³

¹ VTI, Linköping, Sweden

² Department of Biomedical Engineering, Linköping University, Linköping, Sweden

³ Institute of Sport Science, University of Bern, Bern, Switzerland

*christer.ahlstrom@vti.se

Abstract: To successfully get around in traffic it is often necessary to keep track of several relevant targets at the same time. This can be done by combining foveal and peripheral visual information sampling. Especially if no detailed input is needed, for example when confirming the absence of road users, it may be enough to use peripheral vision only. Using a driving simulator with an urban scenery, 35 participants passed three zebra crossings with a) no pedestrians nearby, b) pedestrians standing nearby and c) pedestrians nearby of whom one started walking towards the street. In the last case, all participants foveated the walking person, albeit around one third of the participants already released the throttle before the first glance at the pedestrian. The standing pedestrians were foveated in almost all instances, whereas the roadside nearby the zebra crossing without people nearby was not foveated by around a quarter of the participants. Taken together, the results indicate that peripheral vision may suffice to confirm the absence or presence of pedestrians. With people present, a glance towards them is initiated, likely to check for additional information. Throttle release before foveation is an indication that the walking pedestrian was detected as relevant with peripheral vision.

1. Introduction

When negotiating one's way in traffic it is necessary to attend to relevant targets in time. Several relevant targets can be present in different locations at the same time. By "target" we mean a place or object holding relevant information, including locations that should be checked to confirm the absence or presence of road users.

This information sampling is done predominantly visually. However, to assess whether a driver obtained all necessary information, it is not enough to analyse glance behaviour with an eye tracker. Research both from transportation and sports indicates that peripheral vision plays an important role, not only for the guidance of foveal vision, but also for information acquisition in itself (Vater, Williams, & Hossner, 2020; Wolfe, Sawyer, & Rosenholtz, 2020). It may be more efficient to employ peripheral vision, and in

situations where several targets must be monitored simultaneously, it may even be necessary to employ peripheral vision. For example, when approaching a zebra crossing, it is necessary to assess whether any pedestrians want to cross. Likely, the absence of pedestrians can be confirmed by peripheral vision only, but if pedestrians are present, a foveal glance to estimate their future trajectory will be required, especially if they are or might start moving. Indications for this were found in a field study where zebra crossings often were passed without the driver's glancing to the sides (Kircher & Ahlström, submitted).

Here, we attempt to estimate the usage of peripheral vision by comparing visual sampling strategies across situations that, based on the reasoning above, should either require the usage of foveal vision or not. In addition, we investigate the temporal linkage between action and gaze

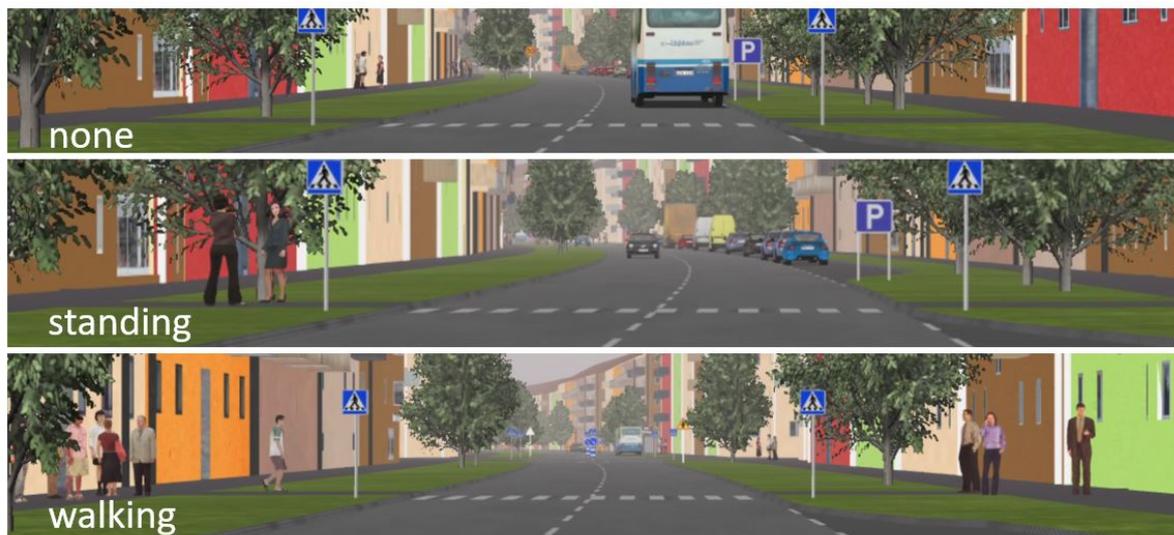


Figure 1. Screenshots of the three zebra crossing situations.

behaviour, again with the goal to deduce a potential employment of peripheral vision.

Method

As part of a study investigating alcohol intoxication and driving, 35 participants drove a route in a simulator in right-hand traffic first sober, then under alcohol influence. Here, we only consider the sober condition. After filling out a background questionnaire, the participants familiarised themselves with the simulator before driving the test route.

1.1 Apparatus

The fixed-base simulator with a visual angle of about 150 degrees was used. Transmission was automatic. A Smart Eye Pro system (Smart Eye AB, Gothenburg, Sweden) with four cameras was used for gaze tracking. The route led through an urban environment (Figure 1). Other road users were present, including pedestrians and cyclists. The route was constructed as a loop where the participants drove one lap, but with two different starting points. The participants had to drive straight on for the whole route and were instructed to drive as usual.

1.2 Selected situations

For the current analysis, we focused on situations for which the use of peripheral vision could be investigated. Specifically, we selected three zebra crossings which were all positioned on a link without being connected to an intersection (Figure 1).

In the “none”-situation, no pedestrians are present in the vicinity of the zebra crossing. A bus leaves a bus stop on the right-hand side of the road and accelerates across the zebra crossing ahead of the participant.

In the “standing”-situation, two pedestrians are standing on the left-hand side of the zebra crossing underneath a tree. They look like they are talking to each other.

In the “walking”-situation, there are groups of pedestrians on both sides of the zebra crossing, somewhat farther away from the road as in the “standing”-situation. During the approach, a pedestrian leaves the group and starts walking towards the zebra crossing and across the street,

timed such that he is on collision course with the driver, requiring a reaction. Note that the “walking”-situation was by no means critical, with plenty of time to slow down to let the pedestrian cross the street.

1.3 Analysis

The analysis focused on glances to the left during the approach to the zebra crossings. The main idea was to check if drivers search for pedestrians at the zebra crossings, even when there are no pedestrians present. If they don't, we hypothesize that they use peripheral vision to determine if there is something worth looking at in that location. In a second step, we also analysed whether drivers started to slow down already before the first glance to the crossing pedestrian, and if so, we hypothesize that peripheral vision is used for action.

The onset of glances to the left was extracted with a velocity-based saccade detection algorithm (threshold = 25°/s, with the following fixation $\geq 3^\circ$ to the left), in combination with an angular-dispersion based algorithm (threshold = 8°). The latter facilitated detection of slow eye movements and smooth pursuits to the left.

In the first analysis, glances within a relevant zone (cf. Kircher & Ahlstrom, 2017) were considered. A chi-square test was used to check for differences between the three zebra crossings in whether drivers glanced to the left or not. In the second analysis, this time window was extended to also include the point from where the zebra crossing became visible. We then assessed the time from throttle release and brake onset, respectively, to the first glance towards the crossing pedestrian. Negative values mean that the driver acted before foveating the pedestrian.

All analyses focused on glances to the left for two reasons. First, the most relevant differences between the three situations were located on the left side of the road. Secondly, glances to the right were difficult to distinguish since the angular difference between the centre of the lane and the right side of the lane is just a few degrees, especially when looked at from a distance.

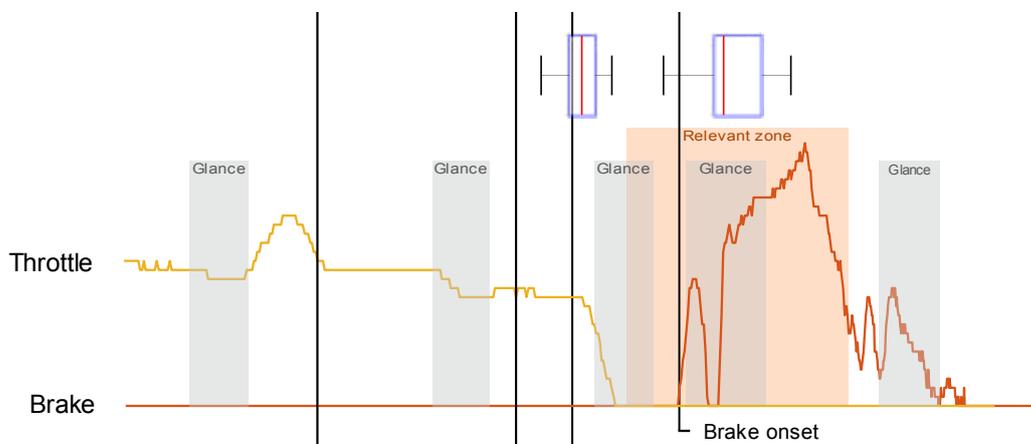


Figure 2. Throttle, brake and glance behaviour for a person releasing the throttle before the pedestrian starts to walk during the “walking”-situation. The boxplot on top illustrates that 35% of the drivers released the throttle before foveating the pedestrian, and 7% started to brake before foveating the walking pedestrian.

2. Results

In the “none”-situation, 27% of the participants did not glance to the left in the relevant zone. In the “standing”-situation, 7% did not glance to the left. In the “walking”-situation all participants glanced left, foveating the pedestrian (chi-square(2)=10.4; p=.006). In that situation, 35% of the participants released the throttle before foveating the pedestrian after he started to move, and 7% even started to brake before foveating the pedestrian.

Figure 2 illustrates one participant’s behaviour in the “walking”-situation, combined with boxplots for throttle release timing and brake onset. The person in the example glances left after the zebra crossing becomes visible (second glance in illustration), but before the pedestrian starts to walk. The throttle is released before the next glance occurs, upon which the brake pedal is pressed.

3. Discussion

Our data indicate that foveal vision is used when pedestrians are present, presumably to identify whether they want to cross the street. However, it also seems like peripheral vision can be enough to confirm the absence of pedestrians in the relevant areas. In the “none”-situation one in four participants did not foveate the left roadside, which on its own could be taken as a sign that they missed checking for potential traffic with priority. However, in combination with the high foveation rates in the other two situations, the finding can also be interpreted such that foveation occurs if information obtained peripherally motivates foveation. In this case, peripheral vision is used to “preview” information and the subsequent fixation to confirm the preview foveally (Vater et al., 2020).

A further indication for peripheral vision being used is the fact that about a third of the participants release the gas pedal before having foveated the pedestrian after he starts walking towards the zebra crossing. This means that they very likely noticed the movement with peripheral vision and prepared themselves for action by releasing the gas before a foveal confirmation of being on collision course with the pedestrian. This then led to further action in the form of braking to let the pedestrian pass. That actions are initiated based on information in peripheral vision is known from sports. For example, high-level martial arts athletes fixate the chest of the opponent and are able to react to the attacking limb (hands or feet) without looking at the limbs (Hausegger, Vater, & Hossner, 2019).

Given the mentioned field study, we would have assumed a higher reliance on peripheral vision to confirm the absence of relevant traffic. It could be speculated that the simulator environment contributes to a more extensive usage of foveal vision. This could be due to the somewhat unnatural movement of cars, pedestrians and cyclists that require more foveation than one would otherwise need to interpret human motion (Blake & Shiffrar, 2007). It could also be connected to the environment being overly simple, such that participants glance around due to understimulation. Therefore, we recommend a follow-up study in the field controlling for pedestrian behaviour.

4. Conclusions

Even though difficult to measure with conventional methods, peripheral vision appears to be an integral part of

visual information acquisition in driving, especially when confirming the absence of relevant traffic. This is important to consider in driver attention monitoring, as an absence of a foveation to a relevant area not necessarily entails that the area was not considered by the driver. We recommend more elaborate studies determining the role and assessment of peripheral vision in driving.

5. Acknowledgments

The study was financed by the Sweden’s Innovation Agency (Vinnova), Grant 2019-05834.

References

- Blake, R., & Shiffrar, M. (2007). Perception of Human Motion. *Annual Review of Psychology*, 58(1), 47-73. doi:10.1146/annurev.psych.57.102904.190152
- Hausegger, T., Vater, C., & Hossner, E.-J. (2019). Peripheral Vision in Martial Arts Experts: The Cost-Dependent Anchoring of Gaze. *Journal of Sport and Exercise Psychology*, 41(3), 137-145. doi:10.1123/jsep.2018-0091
- Kircher, K., & Ahlstrom, C. (2017). Minimum Required Attention: A human-centered approach to driver inattention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 59(3), 471-484. doi:10.1177/0018720816672756
- Kircher, K., & Ahlström, C. (submitted). *Driver attention monitoring and visual sampling from relevant and irrelevant targets*. Paper presented at the Driver Distraction and Inattention, Gothenburg, Sweden.
- Vater, C., Williams, A. M., & Hossner, E. (2020). What do we see out of the corner of our eye? The role of visual pivots and gaze anchors in sport. *International Review of Sport and Exercise Psychology*, 13, 103 - 181.
- Wolfe, B., Sawyer, B. D., & Rosenholtz, R. (2020). Toward a Theory of Visual Information Acquisition in Driving. *Human Factors*, 0018720820939693. doi:10.1177/0018720820939693

Evidence accumulation modeling for the Detection Response Task when combined with the Box Task

Daniel Trommler^{1*}, Tina Morgenstern¹, Cornelia Hollander¹, Ines Karl², Frederik Naujoks², Josef F. Krems¹, Andreas Keinath²

¹ Cognitive & Engineering Psychology, Chemnitz University of Technology, Chemnitz, Germany

² BMW Group, Munich, Germany

*daniel.trommler@psychologie.tu-chemnitz.de

The Box Task (BT) combined with a Detection Response Task (DRT) is a method to assess different dimensions of secondary task demand caused by portable (electronic) devices or in-vehicle systems. This paper presents a comprehensive analysis of the DRT using an evidence accumulation model. The aim was to replicate influences of cognitive load on the model parameter using two secondary tasks. Furthermore, potential practice effects of the BT + DRT were investigated using three baseline drives (i.e., performing the BT + DRT without secondary task engagement) throughout an experimental session. A Wald-distributed model revealed a significant increase in the rate of evidence accumulation across all baseline drives. Additionally, the response omission parameter decreased. In contrast, cognitive task demand resulted in a lower rate of evidence accumulation, a higher response threshold, a higher omission probability and a slightly faster non-decision time. The results suggest a substantial practice effect throughout an experimental session, which highlights the relevance of a sufficient practice duration. Since typical changes in information processing under divided attention were confirmed, it is assumed that the BT + DRT can be used as a valid instrument to assess secondary task demand.

1. Introduction

Minimizing driver distraction is a crucial part when developing in-vehicle information systems (IVIS). To assess the task demand while driving, a relatively new method was developed – the Box Task combined with a Detection Response Task (BT + DRT; Hsieh & Seaman, n. d.). The method is based on the Dimensional Model of Driver Demand (Young et al., 2016), which distinguishes between cognitive and physical (i.e., visual-manual) secondary task demand. Previous studies suggest a high sensitivity of the BT to visual-manual task demand, while the DRT is particularly sensitive to cognitive demand (Morgenstern et al., 2020a; Morgenstern et al., 2020b).

Evidence accumulation modeling describes the process of stimulus detection as an accumulation of stimulus information (evidence) from a starting point to a response threshold (Ratcliff & Van Dongen, 2011). A response to the stimulus is initiated when the amount of accumulated evidence reaches the threshold (Ratcliff & Van Dongen, 2011). Based on reaction times and response accuracies, four key model parameters can be estimated. First, the *rate of evidence accumulation*, which is sensitive to the efficiency in processing stimulus-related information. Second, the height of the *response threshold*. More conservative (i.e., cautious or accurate) responses require more evidence and are reflected in higher response thresholds. Third, the *starting point* of evidence accumulation, which varies from trial-to-trial. And fourth, the *non-decision time*, which summarizes the duration for stimulus encoding as well as motor response execution (Ratcliff & Van Dongen, 2011). Moreover, an *omission parameter* for missing responses was proposed, describing the probability that the stimulus encoding fails or that no evidence can be sampled from an encoded stimulus (Matzke et al., 2017a; Matzke et al., 2017b). Evidence

accumulation models were previously used with the DRT (e.g., Tillman et al., 2017; Castro et al., 2019). A (one-boundary) Wald model revealed that cognitive load was associated with a lower rate of evidence accumulation, a higher response threshold, a faster non-decision time as well as a higher response omission parameter (Castro et al., 2019).

The aim of this paper was to analyze whether the effects of Castro et al.'s (2019) investigation can be replicated for the BT + DRT to ensure its validity. Furthermore, potential practice effects of the BT + DRT were investigated throughout an experimental session.

2. Method

2.1 Participants

In total, 32 participants with a mean age of 27 years ($SD = 6.7$) participated in the study.

2.2 Material

2.2.1 BT + DRT

The Box Task (BT) is designed as a continuous tracking task in which a dynamic box has to be kept within two boundaries using a steering wheel and gas pedal (for a detailed description see Trommler et al., 2021). In parallel, a tactile Detection Response Task (DRT) is used according to the ISO Standard (see ISO 17488, 2016). That means, participants should respond to a vibration stimulus which is placed on their shoulder by pressing a button on the steering wheel. The hit rate and the mean reaction time for hits were captured.

2.2.2 Secondary Tasks

For the conditions with secondary task engagement, two cognitive tasks were used in an easy and difficult version – the *n-back task* (Kirchner, 1958) and *memory scanning task* (MST; Sternberg, 1966). During the *n-back task*, participants were instructed to recall the number at a predetermined position within a series (0-back (easy version) versus 2-back (difficult version)). During the MST, different city names were presented acoustically while the participants had to answer whether the item was included in a memorized set of two (easy version) versus five (difficult version) German city names or not.

2.3 Design and Procedure

The *cognitive secondary task* was a within-subjects factor with five levels: no secondary task, easy and difficult version of secondary task 1 as well as easy and difficult version of secondary task 2. The *duration of practice* of the BT + DRT with 1.5, 3 or 4.5 minutes of practicing was examined as a between-subjects factor. Both independent variables were balanced.

An experimental session started with a practice trial of the BT + DRT depending on the three practice conditions. For each secondary task condition (*n-back task* or MST), there was a test block consisting of two trials (easy and difficult task version). These trials were performed as dual-task conditions, i.e., performing the BT + DRT with secondary task engagement. Before (initial baseline), between (intermediate baseline), and after (final baseline) the two test blocks, three baseline drives (i.e., performing the BT + DRT without secondary task engagement) were conducted. Each trial had a duration of 210 seconds.

3. Results

3.1 Model Estimation

For the model estimation, DMC (Heathcote et al., 2019) was used. A Wald model was employed based on previous findings (e.g., Castro et al., 2019). Thus, a separate rate of evidence accumulation, response threshold, non-decision time and omission probability was calculated for each of the seven trials (three baseline drives and four dual-task conditions). Additionally, the starting point of evidence accumulation varied across trials. The results can be found in Figure 1 to 4 in Appendix A.

3.2 Parameter Tests

For the parameter tests, Bayesian *p*-values were calculated (see Matzke et al., 2015).

3.2.1 Secondary Task Demand

A lower efficiency in processing stimulus-related information under divided attention was confirmed by a significantly decreased rate of evidence accumulation for all dual-task conditions compared to the baseline drives (all $p < .001$). Furthermore, compared to all baseline drives, a significant increase in the response threshold was observed for the dual-task conditions (all $p < .001$). A slight but also significant difference between the baseline and dual-task conditions was found for the non-decision time, with faster encoding and motor execution processes when cognitive load

was present (all $p < .001$). Likewise, the probability of omissions increased significantly for all dual-task conditions compared to the baseline drives (all $p < .001$).

Moreover, all dual-task conditions differed significantly from each other in the rate of evidence accumulation (all $p < .001$), except for the easy and the difficult MST. Additionally, response thresholds were almost equivalent for all dual-task conditions, except for the difficult *n-back task*, which was associated with a significantly lower response threshold than the other dual-task conditions (for the difficult *n-back task* and easy *n-back task*: $p = .021$). The omission parameter revealed significant differences between all dual-task conditions, except for the easy MST and the easy *n-back task* (for difficult MST and easy *n-back task*: $p = .002$; all other: $p < .001$).

3.2.2 Baseline Drives

The rate of evidence accumulation increased significantly between the initial and intermediate baseline drive ($p < .001$) as well as between the intermediate and final baseline drive ($p = .009$). Furthermore, the omission probability decreased significantly between the initial and intermediate baseline drive ($p < .001$). No significant differences were observed in the response threshold and the non-decision time.

4. Conclusion

Previous findings regarding the influence of cognitive load on the DRT were confirmed. Due to limited-capacity attention, the efficiency in information processing decreased and failures in stimulus encoding or evidence sampling increased in the presence of cognitive load. In addition, cognitive load led to a higher response caution as well as a faster non-decision time. These findings underpin the validity of the BT + DRT.

Furthermore, the rate of evidence accumulation increased significantly across the three baseline drives. This suggests a substantial practice effect throughout the experimental session. However, in-depth analyses for the three practice conditions are pending. The results could have implications for a sufficient duration of practicing the BT + DRT.

5. Acknowledgments

This research was supported by the BMW Group.

References

- Castro, S. C., Strayer, D. L., Matzke, D., & Heathcote, A. (2019). Cognitive workload measurement and modeling under divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, 45(6), 826–839. <https://doi.org/10.1037/xhp0000638>
- Heathcote, A., Lin, Y.-S., Reynolds, A., Strickland, L., Gretton, M., & Matzke, D. (2019). Dynamic models of choice. *Behavior Research Methods*, 51(2), 961–985. <https://doi.org/10.3758/s13428-018-1067-y>
- Hsieh, L., & Seaman, S. (n.d.). Evaluation of the Two-Dimensional Secondary Task Demand Assessment Method. Unpublished report, Department of Communication Sciences and Disorders, Wayne State University.
- ISO 17488 (2016). Road vehicles — Transport information and control systems — Detection-response task

(DRT) for assessing attentional effects of cognitive load in driving.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 55(4), 352–358. <https://doi.org/10.1037/h0043688>

Matzke, D., Dolan, C. V., Batchelder, W. H., & Wagenmakers, E.-J. (2015). Bayesian Estimation of Multinomial Processing Tree Models with Heterogeneity in Participants and Items. *Psychometrika*, 80(1), 205–235. <https://doi.org/10.1007/s11336-013-9374-9>

Matzke, D., Hughes, M., Badcock, J. C., Michie, P., & Heathcote, A. (2017a). Failures of cognitive control or attention? The case of stop-signal deficits in schizophrenia. *Attention, Perception, & Psychophysics*, 79(4), 1078–1086. <https://doi.org/10.3758/s13414-017-1287-8>

Matzke, D., Love, J., & Heathcote, A. (2017b). A Bayesian approach for estimating the probability of trigger failures in the stop-signal paradigm. *Behavior Research Methods*, 49(1), 267–281. <https://doi.org/10.3758/s13428-015-0695-8>

Morgenstern, T., Trommler, D., Forster, Y., Naujoks, F., Hergeth, S., Krems, J. F., & Keinath, A. (2020a). Measuring Driver Distraction with the Box Task – A Summary of Two Experimental Studies. In H. Krömker (Ed.), *HCI in Mobility, Transport, and Automotive Systems. Driving Behavior, Urban and Smart Mobility* (Vol. 12213, pp. 51–60). Springer International Publishing. https://doi.org/10.1007/978-3-030-50537-0_5

Morgenstern, T., Wögerbauer, E. M., Naujoks, F., Krems, J. F., & Keinath, A. (2020b). Measuring driver distraction – Evaluation of the box task method as a tool for assessing in-vehicle system demand. *Applied Ergonomics*, 88, 103181. <https://doi.org/10.1016/j.apergo.2020.103181>

Ratcliff, R., & Van Dongen, H. P. A. (2011). Diffusion model for one-choice reaction-time tasks and the cognitive effects of sleep deprivation. *Proceedings of the National Academy of Sciences*, 108(27), 11285–11290. <https://doi.org/10.1073/pnas.1100483108>

Sternberg, S. (1966). High-Speed Scanning in Human Memory. *Science*, 153(3736), 652–654. <https://doi.org/10.1126/science.153.3736.652>

Tillman, G., Strayer, D., Eidels, A., & Heathcote, A. (2017). Modeling cognitive load effects of conversation between a passenger and driver. *Attention, Perception, & Psychophysics*, 79(6), 1795–1803. <https://doi.org/10.3758/s13414-017-1337-2>

Trommler, D., Morgenstern, T., Wögerbauer, E. M., Naujoks, F., Krems, J. F., & Keinath, A. (2021). The box task - A method for assessing in-vehicle system demand. *MethodsX*, 8, 101261. <https://doi.org/10.1016/j.mex.2021.101261>

Young, R., Seaman, S., & Hsieh, L. (2016). The Dimensional Model of Driver Demand: Visual-Manual Tasks. *SAE International Journal of Transportation Safety*, 4(1), 33–71. <https://doi.org/10.4271/2016-01-1423>

Appendix A

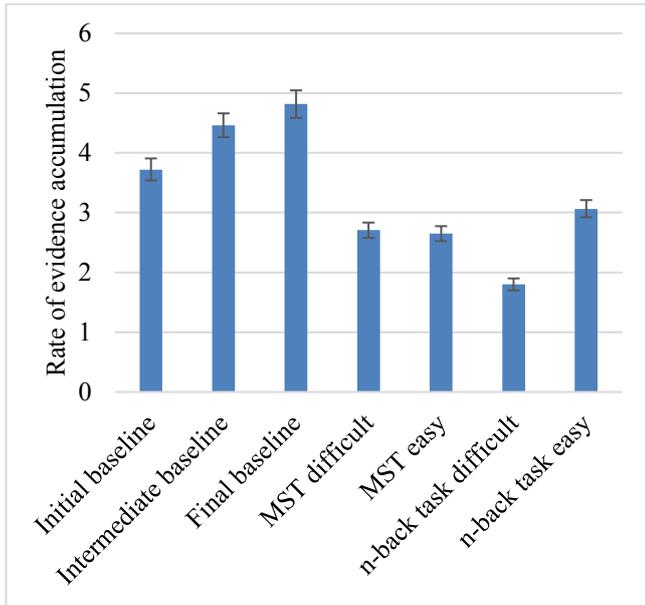


Fig. 1. Rate of evidence accumulation across the task conditions. Error bars represent Bayesian 95% credible intervals.

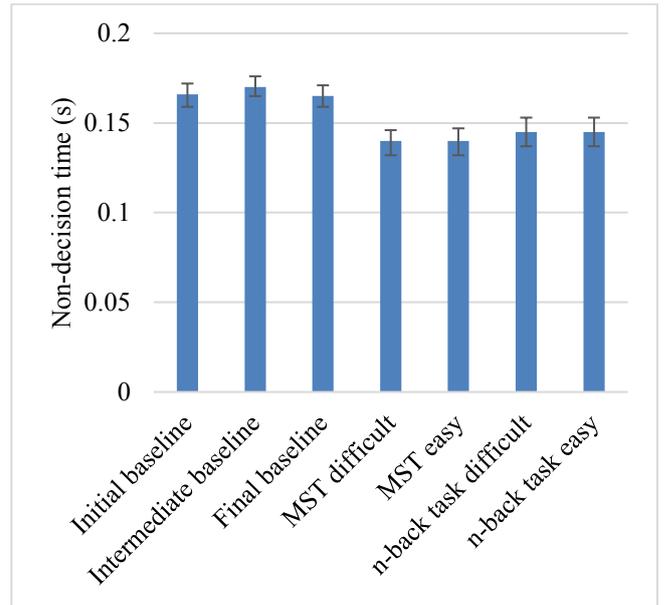


Fig. 3. Non-decision time across the task conditions. Error bars represent Bayesian 95% credible intervals.

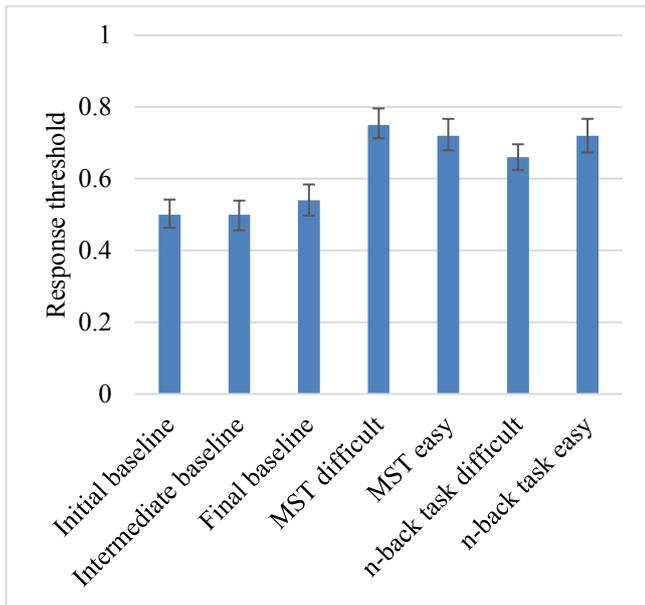


Fig. 2. Response threshold across the task conditions. Error bars represent Bayesian 95% credible intervals.

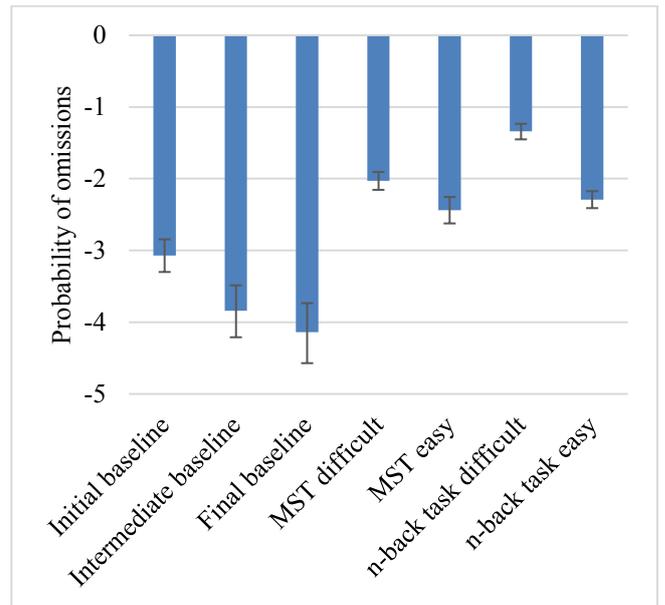


Fig. 4. Probability of omissions across the task conditions. Error bars represent Bayesian 95% credible intervals.

Estimating Driver Fatigue Based on Heart Activity, Respiration Rate

Oren Musicant^{1*}, Bar Richmond-Hacham¹, Assaf Botzer¹

¹ Industrial Engineering & Management, Ariel University, Kiriath Hamada, Ariel, Israel

*musicant.oren@gmail.com

Abstract: Driver inattention contributes to many road accidents, with fatigue playing a prominent role. Hence, driver assistance systems that monitor driver fatigue and intervene (e.g., by alerting the driver) can be beneficial in reducing fatigue-related accidents. This paper reviews the literature on fatigue detection using heart activity (e.g., heart rate variability, heart rate) and respiration indices. We aim to point to the most reliable physiological indices impacted by fatigue and present state-of-the-art driver fatigue detection algorithms and their accuracy. Our review shows that heart rate, systolic blood pressure, and respiration rate are among the most predictive indicators of driver fatigue. Our review also shows that when using valid physiological indices in conjunction with cutting-edge classification techniques, it is possible to differentiate fatigued from non-fatigued drivers with an accuracy sometimes as high as 95%. Our review can assist developers of fatigue detection systems.

1. Introduction

Fatigue, a psychophysiological state characterized by decreased alertness and vigilance, is one of the leading causes of road accidents (Bharadwaj et al., 2021), possibly because fatigued drivers might fail to pay attention to the road and traffic (Wang et al. 2018). Thus, driver assistance systems (DAS) that would monitor drivers' states and alert them about their fatigue might improve driving safety. Accordingly, potential developers of such DAS would benefit from a literature review that outlines the best indicators from a wide range of fatigue indices, e.g., blood pressure (BP), heart rate (HR), heart rate variability (HRV), and respiration rate (RR), as well as the latest classification algorithms, e.g., support vector machine (SVM) and neural network (Jeon et al., 2014; Liang et al., 2008).

This paper presents a review on fatigue detection using heart activity, respiration indices. We aim to point to the most reliable indices impacted by fatigue and present state-of-the-art driver fatigue detection algorithms and their accuracy.

2. Method

The review process involved three stages: (1) identification of relevant literature, (2) a meta-analysis for effect size in two steps, and (3) evaluation of common driver fatigue classification models and their accuracy.

2.1 Search strategy and selection criteria

Relevant literature was identified through Google scholar, IEEE, and ScienceDirect. The search process used a combination of fatigue keywords ("fatigue", "drowsiness") along with physiological indices ("ECG", "HR", "HRV", "Respiration rate"). We also conducted the search with and without driving-related terms ("drivers" and "accidents"). The following inclusion criteria were used: studies reporting accuracy, effect sizes, or statistical significance.

2.2 Sample characteristics

A total of 25 papers were included in this review based on the preceding criteria. The aggregated number of participants was 761. The psychophysiological indices in the retrieved studies were: ECG time-domain indices (RMSSD, RMSDD, pNN50, RRI, HR), ECG frequency-domain indices (Total Power, VLF, LF, HF, LF/HF, LF-Normalized units (NU), HF-Normalized units (NU), asymmetry of the spectrum, median frequency of the power spectrum, BW50), nonlinear-domain indices (SDNN, Entropy, SD1, SD2, SD1/SD2, Hurst exponent, Short scaling exponent) and other parameters, such as systolic blood pressure (SBP), diastolic blood pressure (DBP), inspiration-to-expiration (I: E), and RR. The classification techniques used for driver fatigue detection included Linear Discriminant Analysis (LDA), Multimodal Fusion Recurrent

Neural Network (MFRNN), Neural Network, and Support Vector Machines (SVM).

2.3 General Procedural Framework

Data from each study were analyzed using meta-analytic techniques (see Del Re, 2015; Mikolajewicz & Komarova, 2019). We used a three-stage approach to identify valid physiological indices of driver fatigue. First, we assessed which parameters are significantly affected by fatigue. Next, we examined the pooled effect size (Hedges's g) of the subset of features identified. Last, we presented the accuracy, sensitivity, and specificity of different fatigue detection algorithms.

3. Results

3.1 Parameters Evaluation

Table 1 demonstrates which parameters are significantly affected by fatigue driving (marked in bold text). We only considered parameters that were reported in at least two different studies. Thus, we excluded many of the indices that we listed in sub-section 2.2. The 1st column in Table 1 presents a selection of indicators used to detect driver fatigue, and the 2nd column shows the direction of change in these indicators under fatigue.

The numbers in square brackets in the 2nd column refer to the reference I.D. in Appendix A. The 3rd column in Table 1 shows the 'significant increase to significant decrease ratio'. Finally, the 4th column shows the absolute value of the natural logarithm of the ratio divided by its standard error, with a larger value indicating a greater deviation

from the "no-effect" assumption (see Footnotes a-b, Table 1).

Results reveal that HR ($Ln = 1.78$), LF/HF ($Ln = 1.73$), SDNN ($Ln = 1.66$), VLF ($Ln = 1.62$), and LF ($Ln = 1.60$) are among the physiological indicators that were most markedly and consistently affected by fatigue manipulations, followed by HF ($Ln = 1.35$), I: E ($Ln = 1.29$), SBP ($Ln = 1.17$), and RR ($Ln = 0.80$).

3.2 Overall Effect Size and Heterogeneity

Table 2 shows the effect sizes and heterogeneity of key indices. In driving-simulator studies, the effect sizes ranged between -1.8 to 3.15 and the p-value between 0.000 to 0.084. In real driving studies, the effect sizes ranged between -0.55 to 0.49 and the p-value between 0.289 to 0.424.

Overall, the highest effect estimate was found for HR (Hedge's $g = -1.80$), LF/HF (Hedge's $g = -0.61$), SDNN (Hedge's $g = 2.65$), VLF (Hedge's $g = 3.15$), and LF (Hedge's $g = 0.71$). Regarding the overall direction of change, SDNN, VLF, LF, HF, and I: E had higher values if driving while fatigued, whereas HR, LF/HF, SBP, and RR had lower values.

3.3 Classification Evaluation

Table 3 presents the classification results of fatigue states (fatigued/not fatigued) with different physiological indices, classifiers, and methods to detect fatigue. Accuracy ranged between 58% and 95%; sensitivity ranged between 59% and 95%, and specificity was between 70% and 98%. SVM-based prediction model with entropy data provided

Table 1. Driver fatigue-related changes in HR, Respiratory, and BP indices.

| Indices | Increase Decrease Non-Significant Effect | Increase-to- Decrease Ratio | LN (Increase-to- Decrease) /S.E. |
|---------|--|--------------------------------|--------------------------------------|
| HR | I=0 D=7[1, 4, 6, 12, 17] n.s.=6[13, 22] | 0.33[0.1, 1.12] | 1.78 |
| LF/HF | I=1[8] D=8[2, 3, 6, 8, 10, 14, 15, 17] n.s.=5[6, 13, 22] | 0.36[0.12, 1.14] | 1.73 |
| SDNN | I=5[6, 8] D=0 n.s.=1[22] | 6[0.72, 49.84] | 1.66 |
| VLF | I=5[6, 10, 17] D=0 n.s.=0 | 11[0.61, 198.94] | 1.62 |
| LF | I=10[2, 6, 7, 8, 15, 16] D=3[6, 10, 17] n.s.=4[6, 17] | 2.27[0.83, 6.2] | 1.60 |
| HF | I=9[2, 6, 8, 10, 15, 17] D=3[6, 7, 17] n.s.=6[6, 8, 13] | 1.92[0.75, 4.96] | 1.35 |
| I: E | I=3[20] D=0 n.s.=0 | 7[0.36, 135.52] | 1.29 |
| SBP | I=0 D=3[6, 17] n.s.=2[4, 6] | 0.33[0.05, 2.12] | 1.17 |
| RR | I=1[1] D=3[20] n.s.=1[22] | 0.5[0.09, 2.73] | 0.80 |

Note. SBP: systolic blood pressure; I: E: inspiration-to-expiration; RR: respiration rate. (a) The Increase to decrease ratio used a modified count of the decrease and increase counts: Decrease new = Decrease + 0.5 * non-significant + 0.5 and Increase new = Increase + 0.5 * non-significant + 0.5. This formulation minimizes the differences between the significant increase and significant decrease in the case of a small pull of studies. The 0.5 addition is common for small sample sizes to avoid dividing by zero. (b) The standard error for the Ln (Increase to decrease ratio) is $(1/\text{Increase new} + 1/\text{decrease new})^{0.5}$.

Table 2. Pooled effect size and dispersion of effects.

| Indices | S | K | N | Q ² | I ² (%) | τ ² | Effect size (Hedges's g) | Increase-to- Decrease Ratio | t | P |
|---|---|----|-----|----------------|--------------------|----------------|--------------------------|--------------------------------|-------|--------|
| HR | 5 | 7 | 135 | 4.38 | 0.00 | 0.06 | -1.80 | [-2.4, -1.19] | -7.31 | 0.0003 |
| LF/HF | 7 | 10 | 153 | 1367.74*** | 0.99 | 0.57 | -0.61 | [-1.17, -0.05] | -2.46 | 0.0363 |
| SDNN | 3 | 6 | 59 | 2 | 0.00 | 0.20 | 2.65 | [1.77,3.53] | 7.73 | 0.0006 |
| VLF | 2 | 4 | 80 | 0 | 0.00 | 0.00 | 3.15 | [1.72,4.58] | 7.01 | 0.0059 |
| LF | 4 | 7 | 103 | 0 | 0.00 | 0.00 | 0.71 | [0.54,0.88] | 10.46 | 0.0000 |
| HF | 4 | 10 | 103 | 0.17 | 0.00 | 0.00 | 0.46 | [-0.07,0.99] | 1.98 | 0.0787 |
| I: E | 1 | 3 | 30 | 1490.63*** | 1.00 | 0.38 | 1.34 | [-0.19,2.87] | 3.76 | 0.0640 |
| SBP | 3 | 5 | 115 | 5.81 | 0.31 | 0.39 | -1.38 | [-2.44, -0.32] | -3.63 | 0.0222 |
| RR | 2 | 4 | 39 | 35.15*** | 0.91 | 0.96 | -1.35 | [-3.03,0.34] | -2.55 | 0.0841 |
| Data from real driving studies - only indices from at least two different studies | | | | | | | | | | |
| LF/HF | 3 | 4 | 39 | 548.58*** | 0.99 | 1.34 | -0.55 | [-2.42,1.33] | -0.92 | 0.4242 |
| SDNN | 2 | 3 | 19 | 0.04 | 0.00 | 0.00 | 0.49 | [-0.98,1.96] | 1.43 | 0.2886 |

Note. SBP: systolic blood pressure; I: E: inspiration-to-expiration; RR: respiration rate. Q²: heterogeneity between effect sizes; I²: the proportion of heterogeneity; τ²: The absolute value of true variance across studies; S: The number of studies meeting the inclusion criteria for this meta-analysis; K: The number of estimates (sometimes more than one per study); N: Accumulated sample size. Bold text indicates significant Hedge's g values (P < 0.05); *** p<0.001.

the best accuracy compared to the other classification models, with sensitivity, specificity, and accuracy of 95%.

The neural network model applied to the spectral image of the PSD data and MFRNN model with HR data were also found to have good accuracy (90% and 92% accuracy, respectively). On the other hand, the LDA model with HR data

provided a relatively low classification accuracy (64%). Finally, using the LDA model with HR and respiratory indicators combined yielded the lowest accuracy (58%) but the highest specificity (98%).

4. Discussion and Conclusions

The objective of this review was to assess the feasibility of estimating driver fatigue using heart

Table 3. Fatigue Classification Results

| Study | Method | N | Indices | Model | Classification Results (%) |
|-----------------------|--------------------------------|----|--|--------------------|---------------------------------------|
| Salvati et al. (2021) | PERCOLS | 3 | 0.0017Hz,0.0035Hz,0.0053Hz | Analytical Formula | Acc =63 |
| Vicente et al. (2016) | Video – expert evaluation | 30 | Respiratory frequency min (NU), Respiratory frequency min, LF (NU) median absolute deviation, LF median absolute deviation, LF (NU) min, LF min, HRI min | LDA | Acc =58 PPV=96 Se =59 Sp =98 |
| Awais et al. (2017) | Video - expert evaluation | 22 | LF(NU), HF (NU), LF/HF (NU), LF(NU), HF (NU), LF/HF, VLF (NU) | SVM | Acc =70 Se =69 Sp =70 |
| Li & Chung (2013) | PERCOLS | 4 | Entropy | SVM | Acc =95 Se =95 Sp =95 |
| Li & Chung (2013) | PERCOLS | 4 | LF/HF | SVM | Acc =69 Se =62 Sp =75 |
| Ahn et al. (2016) | Well-rested vs. sleep-deprived | 11 | HR | LDA | Acc =64 |
| Patel et al. (2011) | Video - expert evaluation | 12 | Spectral image of the PSD | Neural network | Acc =90 |
| Du et al. (2021) | KSS | 20 | HR | MFRNN | Acc =92 |

Note. LDA: linear discriminant analysis; SVM: support vector machine; PERCLOS: a measure of eye-closure; MFRNN: multimodal fusion recurrent neural network; KSS: Karolinska sleepiness scale; Acc: accuracy; Se: sensitivity; Sp: specificity; PPV: positive predictive value.

rate and respiratory indices. We evaluated which physiological indicators are most useful for detecting driver fatigue and surveyed the utility of different data reduction methods and classification algorithms.

Results showed that HR, VLF, SDNN, LF, and LF/HF are promising indicators of driver fatigue, supporting previous findings (Chen et al., 2021). These findings are based on studies that utilized various evaluation methods (e.g., real-world observations and simulator experiments). It is important to note that the effect sizes and p-values obtained from real-driving studies were smaller than those of simulator studies. This is not surprising given the lack of research examining driver fatigue in real-road driving. More real-road driving studies are needed.

Accuracy, sensitivity, and specificity in the studies reviewed ranged between 58%-95%, 59%-95%, and 70% -98%, respectively. These values indicate that heart activity indices might have a strong potential for detecting fatigue. However, note that our review consists of studies that analyzed various data sets using different classification algorithms, making it difficult to compare their results. Therefore, future empirical studies should test their performance on the same data sets to provide reliable comparisons between classification models. We also recommend that future research in driver fatigue detection investigate the utility of personalized detection algorithms.

Despite the need for more empirical data on driver fatigue detection and for different approaches for analyzing this data (e.g., running different classifiers on the same data set), the findings of our review still offer valuable insights for both DAS developers and investors regarding DAS development considerations, implementation, and evaluation. Further, this study has potential applications for road safety researchers interested in biosignal-based driver fatigue detection.

5. Acknowledgments

The research was funded by Renault Nissan Mitsubishi Alliance, The Israeli branch.

References

- Ahn, S., Nguyen, T., Jang, H., Kim, J. G., & Jun, S. C. (2016). Exploring Neuro-Physiological Correlates of Drivers' Mental Fatigue Caused by Sleep Deprivation Using Simultaneous EEG, ECG, and fNIRS Data. *Frontiers in Human Neuroscience*, 10. <https://doi.org/10.3389/fnhum.2016.00219>
- Argyle, E. M., Marinescu, A., Wilson, M. L., Lawson, G., & Sharples, S. (2021). Physiological indicators of task demand, fatigue, and cognition in future digital manufacturing environments. *International Journal of Human-Computer Studies*, 145, 102522. <https://doi.org/10.1016/j.trf.2018.07.011>
- Awais, M., Badruddin, N., & Drieberg, M. (2017). A Hybrid Approach to Detect Driver Drowsiness Utilizing Physiological Signals to Improve System Performance and Wearability. In *Sensors* (Vol. 17, Issue 9). <https://doi.org/10.3390/s17091991>
- Balasubramanian, V., & Bhardwaj, R. (2018). Can cECG be an unobtrusive surrogate to determine cognitive state of driver? *Transportation Research Part F Traffic Psychology and Behaviour*, 58, 797–806. <https://doi.org/10.1016/j.trf.2018.07.011>
- Beanland, V., Fitzharris, M., Young, K. L., & Lenné, M. G. (2013). Driver inattention and driver distraction in serious casualty crashes: Data from the Australian National Crash In-depth Study. *Accident Analysis & Prevention*, 54, 99–107. <https://doi.org/https://doi.org/10.1016/j.aap.2012.12.043>
- Bharadwaj, N., Edara, P., & Sun, C. (2021). Sleep disorders and risk of traffic crashes: A naturalistic driving study analysis. *Safety science*, 140, 105295. <https://doi.org/https://doi.org/10.1016/j.compbiomed.2021.104413>
- Chen, S., Xu, K., Yao, X., Zhu, S., Zhang, B., Zhou, H., Guo, X., & Zhao, B. (2021). Psychophysiological data-driven multi-feature information fusion and recognition of miner fatigue in high-altitude and cold areas. *Computers in Biology and Medicine*, 133, 104413. <https://doi.org/https://doi.org/10.1016/j.aap.2017.11.038>
- Jacobé de Naurois, C., Bourdin, C., Stratulat, A., Diaz, E., & Vercher, J. L. (2019). Detection and prediction of driver drowsiness using artificial neural network models. *Accident; analysis and prevention*, 126, 95–104. <https://doi.org/10.1016/j.aap.2017.11.038>
- Del Re, A. C. (2015). A Practical Tutorial on Conducting Meta-Analysis in R. *The Quantitative Methods for Psychology*, 11, 37–50. <https://doi.org/10.20982/tqmp.11.1.p037>
- Du, G., Li, T., Li, C., Liu, P. X., & Li, D. (2020). Vision-Based Fatigue Driving Recognition Method Integrating Heart Rate and Facial Features. *IEEE Transactions on Intelligent Transportation Systems*, 22(5), 3089–3100. <https://doi.org/10.1109/TITS.2020.2979527>
- Heikoo, D. D., de Winter, J. C. F., van Arem, B., & Stanton, N. A. (2019). Acclimatizing to automation: Driver workload and stress during partially automated car following in real traffic. *Transportation Research Part F: Traffic Psychology and Behaviour*, 65, 503–517. <https://doi.org/https://doi.org/10.1016/j.trf.2019.07.024>
- Huang, Y. C., Li, L. P., Liu, Z. G., Zhu, H. Y., & Zhu, L. (2019). Assessment of Urban Railway Transit Driver Workload and Fatigue under Real Working Conditions. *Transportation research record*, 2673(11), 891–900. <https://doi.org/10.1177/0361198119826071>
- Jeon, S. Y., Noh, S. J., & Whang, M. C. (2014). Classification of mental fatigue induced by sympathetic or parasympathetic hyperactivity. *대한인간공학학회 학술대회논문집*, 107–112.
- Jung, S.-J., Shin, H.-S., & Chung, W.-Y. (2014). Driver Fatigue and Drowsiness Monitoring System with Embedded Electrocardiogram Sensor on Steering Wheel. *Intelligent Transport Systems, IET*, 8, 43–50. <https://doi.org/10.1049/iet-its.2012.0032>
- Khushaba, R., Kodagoda, S., Lal, S., & Dissanayake, G. (2013). Uncorrelated fuzzy neighborhood preserving analysis based feature projection for driver drowsiness recognition. *Fuzzy Sets and Systems*, 221, 90–111. <https://doi.org/10.1016/j.fss.2012.12.003>
- Ebrahimian, S., Nahvi, A., Bakhoda, H., Homayounfard, A., & Tashakori, M. (2020). Evaluation of driver drowsiness using respiration analysis by thermal imaging on a driving simulator. *Multimedia Tools and Applications*, 79, 17793–17815. <https://doi.org/10.1007/s11042-020-08696-x>
- LAL, S., & Craig, A. (2002). Driver fatigue: Electroencephalography and psychological assessment. *Psychophysiology*, 39, 313–321. <https://doi.org/10.1017/S0048577201393095>
- Lee, S., Kim, M., Jung, H., Kwon, D., Choi, S., & You, H. (2020). Effects of a Motion Seat System on Driver's Passive Task-Related Fatigue: An On-Road Driving Study.

- Sensors*, 20, 2688.
<https://doi.org/10.3390/s20092688>
- Li, G., & Chung, W.-Y. (2013). Detection of Driver Drowsiness Using Wavelet Analysis of Heart Rate Variability and a Support Vector Machine Classifier. In *Sensors* (Vol. 13, Issue 12).
<https://doi.org/10.3390/s131216494>
- Liang, W. C., Yuan, J., Sun, D. C., & Lin, M. H. (2007). Variation in physiological parameters before and after an indoor simulated driving task: effect of exercise break. In *International Conference on Gerontechnology and Service Management (ICGTSM)*.
- Liang, W. C., Yuan, J., Sun, D. C., & Lin, M. H. (2008). Changes in physiological parameters induced by simulated driving tasks: morning vs. afternoon session. *Journal of the Chinese Institute of Industrial Engineers*, 25(6), 457–471.
<https://doi.org/10.1080/10170660809509108>
- MacLean, A. W., Davies, D. R. T., & Thiele, K. (2003). The hazards and prevention of driving while sleepy. *Sleep Medicine Reviews*, 7(6), 507–521.
[https://doi.org/https://doi.org/10.1016/S1087-0792\(03\)90004-9](https://doi.org/https://doi.org/10.1016/S1087-0792(03)90004-9)
- Mikolajewicz, N., & Komarova, S. V. (2019). Meta-Analytic Methodology for Basic Research: A Practical Guide. *Frontiers in Physiology*, 10.
<https://doi.org/10.3389/fphys.2019.00203>
- Patel, M., Lal, S. K. L., Kavanagh, D., & Rossiter, P. (2011). Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert Systems with Applications*, 38(6), 7235–7242.
<https://doi.org/https://doi.org/10.1016/j.eswa.2010.12.028>
- Bhardwaj, R., & Balasubramanian, V. (2019). Viability of Cardiac Parameters Measured Unobtrusively Using Capacitive Coupled Electrocardiography (cECG) to Estimate Driver Performance. *IEEE Sensors Journal*, 19(11), 4321–4330.
<https://doi.org/10.1109/JSEN.2019.2898450>
- Traffic Injury Prevention*, 8(3), 290–299.
<https://doi.org/10.1080/15389580701257842>
- Rodriguez-Ibañez, N., García-Gonzalez, M. A., De La Cruz, M. A. F., Fernández-Chimeno, M., & Ramos-Castro, J. (2012). Changes in heart rate variability indexes due to drowsiness in professional drivers measured in a real environment. *Computing in Cardiology*, 39, 913–916.
- Salvati, L., D'amore, M., Fiorentino, A., Pellegrino, A., Sena, P., & Villecco, F. (2021). On-road detection of driver fatigue and drowsiness during medium-distance journeys. *Entropy*, 23(2), 1–12.
<https://doi.org/10.3390/e23020135>
- Siddiqui, H. U. R., Saleem, A. A., Brown, R., Bademci, B., Lee, E., Rustam, F., & Dudley, S. (2021). Non-invasive driver drowsiness detection system. *Sensors*, 21(14).
<https://doi.org/10.3390/s21144833>
- Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis & Prevention*, 35(3), 381–391.
[https://doi.org/https://doi.org/10.1016/S0001-4575\(02\)00014-3](https://doi.org/https://doi.org/10.1016/S0001-4575(02)00014-3)
- Vicente, J., Laguna, P., Bartra, A., & Bailón, R. (2016). Drowsiness detection using heart rate variability. *Medical & Biological Engineering & Computing*, 54(6), 927–937.
<https://doi.org/10.1007/s11517-015-1448-7>
- Viswanathan, M., Zhang, Z. X., Tian, X. W., & Lim, J. S. (2011). A fatigue detection algorithm by heart rate variability based on a neuro-fuzzy network. *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication, ICUIMC 2011*, 1–6.
<https://doi.org/10.1145/1968613.1968712>
- Zhao, C., Zhao, M., Liu, J., & Zheng, C. (2012). Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator. *Accident Analysis and Prevention*, 45, 83–90.
<https://doi.org/10.1016/j.aap.2011.11.019>
- Wang, Y., Zhu, H., Zhou, F., & Zhao, Y. (2018). How driving duration influences drivers' visual behaviors and fatigue awareness: A naturalistic truck driving test Study. *Ethiopian Journal of Health Development*, 32(1).

Appendix A

Table A1. A list of reviewed studies according to their domain, methodological characteristics and reported measures

| Ref I.D. Ref | Domain | Method | Indices | Subjects | Sig | Estimates | Accuracy |
|--------------------------------------|----------------|--------------------------|---|----------|-----|-----------|----------|
| 1 Argyle et al. (2021) | Psychology | Laboratory | HR, RR | 34 | 2 | 2 | 0 |
| 2 Balasubramanian & Bhardwaj (2018) | Driving | Simulator | LF, HF, LF/HF | 35 | 3 | 3 | 0 |
| 3 Lee et al. (2020). | Driving | Real Driving | LF/HF | 20 | 1 | 1 | 0 |
| 4 Lal & Craig (2002) | Driving | Simulator | DBP, SBP, HR | 35 | 3 | 3 | 0 |
| 5 Salvati et al., (2021) | Driving | Real Driving | 0.0017Hz, 0.0035Hz, 0.0053Hz | 3 | 0 | 0 | 1 |
| 6 Liang et al. (2008) | Driving | Simulator | SBP, DBP, HR, SDNN, HF, LF, VLF, LF/HF | 40 | 30 | 30 | 0 |
| 7 Zhao & Zheng (2012) | Driving | Simulator | Entropy, HF, LF | 13 | 3 | 3 | 0 |
| 8 Rodriguez-Ibañez et al. (2012) | Driving | Real Driving | RMSDD, RRI, SDNN, HF, LF, LF/HF, and several spectrum indices | 10 | 22 | 22 | 0 |
| 9 Vicente et al. (2016) | Driving | Simulator & Real Driving | 7 indices based on Respiratory frequency, HR, and LF region. | 30 | 7 | 7 | 1 |
| 10 Awais & Driberg (2017) | Driving | Simulator | VLF, LF, HF, LF/HF, VLF (both normalized and non-normalized values) | 22 | 4 | 4 | 1 |
| 11 Li & Chung (2013) | Driving | Simulator | LF/HF, Entropy | 4 | 0 | 0 | 2 |
| 12 Ahn et al. (2016) | Driving | Simulator | HR | 11 | 1 | 1 | 1 |
| 13 Jung, Shin & Chung (2014) | Driving | Field test | HF, HR, LF/HF, pNN50, RMSSD, RRI | 2 | 15 | 15 | 0 |
| 14 Patel et al. (2011) | Driving | Simulator | Spectral image of the PSD, LF/HF | 12 | 1 | 1 | 1 |
| 15 Bhardwaj & Balasubramanian (2019) | Driving | Simulator | LF, HF, LF/HF, SD1, SD2, SD1/SD2 | 20 | 6 | 6 | 0 |
| 16 Huang et al. (2019) | Driving | Real Driving | LF | 6 | 3 | 3 | 0 |
| 17 Liang et al. (2007) | Driving | Simulator | SBP, HF, HR, LF, LF/HF, RMMSD, RRI, TP, VLF | 40 | 13 | 13 | 0 |
| 18 Du et al. (2020) | Driving | Simulator | HR | 20 | 0 | 0 | 1 |
| 19 Viswanathan et al. (2011) | Driving | Simulator | SDNN, RMSSD, VLF, LF, HF, LF/HF | 20 | 0 | 0 | 1 |
| 20 Kiashari et al. (2020) | Driving | Simulator | RR, SD RR, I: E, SD I: E, RR, SD RR, I: E, SD I: E | 30 | 6 | 6 | 10 |
| 21 Siddiqui et al. (2021) | Driving | Real Driving | RR | 40 | 0 | 0 | 1 |
| 22 Heikoop et al. (2019) | Driving | Real Driving | HR, LF/HF, SDNN, RR | 9 | 4 | 4 | 0 |
| 23 Babaieian & Mozumdar (2019) | Control Center | Simulator | 24 features in frequency domains from the LF and HF regions | 25 | 0 | 0 | 16 |
| 24 Khushaba et al. (2013) | Driving | Simulator | 23 indices | 31 | 0 | 0 | 3 |
| 25 de Naurois et al. (2019) | Driving | Simulator | RRI, SDNN, CV RMSSD, RR, SDRR, CVRR, RMSSD RR, RRI, SDNN, CVRRI | 8 | 0 | 0 | 3 |

Note. SBP: systolic blood pressure; DBP: diastolic blood pressure I: E: inspiration-to-expiration; RR: respiration rate; HR: heart rate; Sig: the number of significant reports; Estimates: the number of effect size reports; Accuracy: the number of accuracy estimates.

What leads drivers to illicitly nap during conditionally automated driving?

Johanna Wörle*, Barbara Metz

Würzburger Institut für Verkehrswissenschaften (WIVW) GmbH, Robert-Bosch-Straße 4, Veitshöchheim, Germany

*woerle@wivw.de

Abstract: Automation misuse can cause traffic hazards when drivers over-rely on automation and use it in ways that are not intended by the designers of the system. Automation abuse refers to designers of automation designing systems without regard to the consequences for human performance. In a driving simulator study, half of the participants were observed sleeping at least once during six drives with a conditionally automated driving (CAD) system. Sleep is an illicit driver behaviour in CAD because drivers must be available to take over vehicle control at system boundaries. However, sleep was not only observed in the driving simulator environment, but nearly half of the participants indicated that they intend to sleep during CAD in real life. CAD usage, gaze behaviour, subjective evaluation of CAD, trust and mental model of CAD were compared for participants who indicated they intended to sleep in CAD and participants who indicated no intention to sleep. The majority of participants understood that sleep is an illicit driver behaviour in CAD. Participants with the intention to sleep used the simulated CAD more and they reported higher comfort levels during CAD usage and perceived takeover situations as safer. Semi-structures interviews after the last drive indicated that drivers would sleep during CAD once they had some experience with the system. The results suggest that drivers, after gaining experience with CAD, become complacent and sleep during CAD even though they know that it can potentially lead to dangerous situations. Sleep during CAD is both automation misuse and automation abuse. Driver monitoring systems for CAD must detect and prevent sleep in drivers.

1. Introduction

When human operators of automated systems “rely uncritically on automation without recognizing its limits”, they may use it in ways not intended by the designers of the automation. On the other hand, designers of automation might design systems without considering the consequences for human behaviour (Parasuraman & Riley, 1997). In our driving simulator study on “naturalistic” usage of a conditionally automated driving (CAD) system, we observed that 14 of 30 participants napped at least once during six drives. All participants were instructed to remain “sufficiently alert” to be able to resume control of the vehicle at any time during the drives. In CAD (level 3 according to SAE, 2021), the driver must respond with a short notice to a request to intervene at system limits or system failures. In our study, participants did not receive a warning when they were classified as “unavailable” to take back control as it is requested for CAD systems (UNECE, 2021). A request to intervene was only issued when drivers reached sleep stage N2 (stable sleep according to AASM, 2017). After completing six drives with the CAD in the simulator, drivers were asked about their intention to use CAD in real life. Half of the sample said they would sleep during CAD in real life.

Over-trust, high workload and a low (perceived) risk are associated with the misuse of automation (Parasuraman & Riley, 1997). In a naturalistic driving study, misuse of automation due to overconfidence in the system’s capabilities led to 57% of all safety-critical events (Kim, Song, & Doerzaph, 2020). A wrong understanding of the system or wrong ‘mental model’ is one factor that leads to over-trust and over-confidence in the system (Abraham, Seppelt, Mehler, & Reimer, 2017; Seppelt & Victor, 2020). Studies show that partially and conditionally automated driving contribute to the development of drowsiness (Neubauer,

Matthews, & Saxby, 2014; Schömig, Hargutt, Neukum, Petermann-Stock, & Othersen, 2015; Vogelpohl, Kühn, Hummel, & Vollrath, 2019). The observed instances of sleep during our study could be partially explained by the simulator environment. However, half of the participants stated that they would sleep during CAD in real life. The aim of the presented study was to understand why drivers intend to sleep during CAD.

2. Method

The principle objective of the driving simulator study was to investigate behavioural adaptation to a CAD. Participants were invited to take part in six driving sessions in a high-fidelity driving simulator. During each drive, they could use a CAD system for motorways. Participants were instructed with the wording of §1b of the German Road Traffic Act, which specifies the responsibilities of the driver when using CAD:

“[...] the driver may divert his attention from other traffic and control of the vehicle; he must, however, remain sufficiently alert that he can comply with the obligation [to retake control in response to a request to intervene]”

During each drive, participants experienced system boundaries and requests to intervene with a takeover time budget of 15s. Two of the driving sessions, the Baseline drive and the Sleepy drive (see Table 1), were designed with the aim of investigating the effects of fatigue and sleepiness. The EEG was measured during both drives. The Sleepy drive was scheduled at 6 a.m. and participants were sleep deprived to promote sleepiness during the drive. The driving environment was designed to be monotonous in both drives, with low traffic volume and fog to limit visibility. Sleep stages were coded according to the American Academy of Sleep Medicine standard (AASM, 2017) based on EEG. Eye-tracking parameters were measured using a SmartEye® four-

camera system. Driving and system parameters were recorded using Silab®. For a more detailed description of the study design and procedure, please refer to Metz et al., 2021.

Table 1 Overview study sessions

| Session | Group 1 | Group 2 |
|---------|-----------------------|-----------------------|
| 1 | Short drive | Short drive |
| 2 | Short drive | Short drive |
| 3 | Baseline drive | Sleepy drive |
| 4 | Short drive | Short drive |
| 5 | Sleepy drive | Baseline drive |
| 6 | Short drive | Short drive |

2.1 Post-drive questionnaire and interviews

After each drive, participants completed a short version of the L3Pilot common questionnaire (Metz et al., 2020), which included questions on trust and attitudes towards CAD, willingness to use and mobility-related questions. The scale ranged from 1 (strongly disagree) to 5 (strongly agree). The questions on the mental model of CAD were added specifically for this study.

Although the Sleepy drive was designed to induce sleepiness, we did not expect participants to sleep. After observing participants falling asleep, we designed an interview guideline on intention to sleep during AD and conducted post-drive interviews with a subsample. The interview guideline included the following questions, among others:

1. Did you sleep during the study when the automated driving system was active? Did you sleep intentionally?
2. If you could use such a system in real life, would you sleep when it was active?
3. Is it possible to respond appropriately to a request to intervene when you are asleep?

2.2 Sample

N = 31 participants (13 female, mean age = 37, SD = 12) took part in the study. The interviews were conducted with a subsample of 22 participants (7 female, mean age = 41, SD = 12). All participants held a valid driving license and had completed an extensive driving simulator training.

2.3 Data analysis

The data of one participant were excluded from the analyses due to data loss, resulting in a final sample of N = 30 participants. A Multivariate Analysis of Variance (MANOVA) was performed to compare the effects of relevant behavioural measures and questionnaire responses between participants with the intention to sleep during CAD and participants without the intention to sleep during CAD. The dependent variables were:

- **System usage (%)**: proportion of time the system was activated (measured with Silab)
- **NDRA (%)**: proportion of driving with CAD which was spend on non-driving related activities (coded by the experimenter throughout the drives)
- **PRC**: Percentage Road Center, proportion of time the participant’s gaze was directed to the windshield (measured with SmartEye®)

- **PerCLOS**: Percentage of eyelid closure, an eye-tracking based measure of driver drowsiness (Dinges & Grace, 1998), measured with SmartEye®)
- **Willingness to use**: “I would use this system if it was in my car.”
- **Perceived safety**: “I felt safe when driving with the system active.”
- **Workload**: “Driving with this system was demanding.”
- **Trust**: “I trust the system to drive.”
- **Comfort**: “Driving with the system active was comfortable.”
- **Increased drowsiness**: “Driving with the function on long journeys would make me tired.”
- **Safety during takeover**: “During the takeover I always felt safe.”

3. Results

14/30 participants experienced EEG-verified sleep at least once (Observed Behaviour). In the questionnaire after the sixth driving session, 15/30 participants stated that they would sleep at least very infrequently if they had CAD in their car (Behavioural intention). In the same questionnaire, participants were asked about their mental model of CAD. They had to indicate if a statement was correct (Yes) or incorrect (No) or if they were not sure (I don’t know). 2/30 participants stated that sleeping is allowed in CAD and three participants were not sure if it is allowed (Mental model, see Table 2).

Table 2 Overview of observed behaviour during the study, behavioural intention in real usage and mental model of CAD

| Observed behaviour | Behavioural intention | Mental model (Sleep allowed) | | |
|--------------------|-----------------------|------------------------------|----|--------------|
| | | Yes | No | I don’t know |
| Sleep | Yes | 2 | 6 | 3 |
| | No | 0 | 3 | 0 |
| No sleep | Yes | 0 | 4 | 0 |
| | No | 0 | 12 | 0 |

3.1 Behavioural data and questionnaire data

The MANOVA revealed significant effects of system usage, NDRA engagement, willingness to use, comfort and perceived safety during takeover on the behavioural intention to sleep (for an overview of statistical figures, see Table 3).

Table 3 Means, Standard Deviations and One-Way ANOVA of observed behaviour and questionnaire answers based on the Behavioural intention to sleep

| Measure | Behavioural intention | | | | F(1, 27) | p |
|------------------------|-----------------------|-------------|-------------|-------------|--------------|-------------|
| | Yes | | No | | | |
| | M | SD | M | SD | | |
| System usage (%) | 0.93 | 0.03 | 0.87 | 0.11 | 4.267 | .049 |
| NDRAs (%) | 0.79 | 0.14 | 0.56 | 0.35 | 6.210 | .019 |
| PRC | 0.23 | 0.10 | 0.37 | 0.24 | 4.095 | .053 |
| PerCLOS | 0.14 | 0.12 | 0.18 | 0.24 | 0.543 | .467 |
| Willingness to use | 4.53 | 0.52 | 3.71 | 1.20 | 5.806 | .023 |
| Perceived safety | 4.07 | 0.70 | 3.50 | 1.10 | 2.799 | .106 |
| Workload | 1.60 | 0.51 | 1.93 | 1.00 | 1.277 | .268 |
| Trust | 3.40 | 1.55 | 3.71 | 0.82 | 0.455 | .506 |
| Comfort | 4.53 | 0.64 | 3.71 | 1.10 | 6.370 | .018 |
| Increased drowsiness | 3.40 | 1.24 | 3.50 | 2.03 | 0.026 | .873 |
| Safety during takeover | 4.60 | 0.51 | 3.64 | 1.15 | 8.605 | .007 |

3.2 Interview data

When interviewed after the drive, 7/22 drivers stated that they would sleep when using a CAD system. They stated that time would pass more quickly and they would catch up on sleep. Some participants indicated that they would only sleep under certain conditions, for example only in low traffic scenarios or only on familiar routes. Some participants indicated that they would observe the system first and if it worked as intended, they would feel safe enough to sleep. In general, participants who intended to sleep indicated that it would make their journey easier. Only one participant believed that it is possible to respond appropriately to a request to intervene after sleep.

4. Discussion

Sleep is an illicit behaviour during CAD and it can lead to dangerous situations if drivers are not able to take over vehicle control at system boundaries. Despite repeated experience with system boundaries, half of the participants of our driving simulator study stated their intention to sleep during CAD. The majority of the sample was aware that sleeping is not allowed in CAD. Thus, a wrong system understanding was not the reason for participants' intention to sleep. Contrary to Parasuraman and Riley's (1997) definition of automation misuse as being associated to over-trust, we found no relationship between trust and intention to sleep. Participants who were generally more willing to use CAD and used it more frequently during the study indicated their intention to sleep in CAD. This was also reflected in interview statements that participants found it useful and comfortable to sleep while travelling. Increased sleepiness due to automation or the objective drowsiness during CAD

use, as measured with PerCLOS, did not influence the intention to sleep in CAD. The perceived safety of driving with CAD did not have an effect, but the perceived safety during takeovers had an effect on the intention to sleep. One explanation for this could be that drivers who experience takeover situations as safe might believe that they can handle these situations safely after waking up from sleep. It seems that after drivers gain experience with the system, they become complacent (Parasuraman & Manzey, 2010). Although they are aware of system boundaries, they develop the false feeling that "everything is fine" when in fact, sleep can lead to hazards in takeover situations (Wörle, Metz, Othersen & Baumann, 2020). However, it has to be taken into account that in our study, although we used EEG to monitor driver state and detect sleep in drivers, we did not warn participants before they fell asleep. That way, the CAD system enabled drivers to sleep and did not prevent them from falling asleep. Drivers sleeping during CAD in our study and drivers' intention to sleep is an abuse of automation. CAD enables drivers to retrieve from the driving task and therefore increases the risk for sleep.

5. Conclusions

Despite knowing that it is not allowed, drivers might become complacent and sleep when using a CAD system. Sophisticated driver monitoring systems should be implemented not only to detect drowsiness, but also to prevent drivers from falling asleep. If a driver falls asleep, a minimal risk maneuver should be initiated to ensure safety.

6. Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101006664. The sole responsibility of this publication lies with the authors. The authors would like to thank all partners within Hi-Drive for their cooperation and valuable contribution.

References

- AASM. (2017). *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Darien, IL: American Academy of Sleep Medicine.
- Abraham, H., Seppelt, B., Mehler, B., & Reimer, B. (2017). *What's in a name: Vehicle technology branding & consumer expectations for automation*. Paper presented at the Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications.
- Dinges, D. F., & Grace, R. (1998). *PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance*. Retrieved from Washington, D.C.:
- Kim, H., Song, M., & Doerzaph, Z. (2020). *Real-World Use of Partially Automated Driving Systems and Driver Impressions*. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.
- Metz, B., Rösener, C., Louw, T., Aittoniemi, E., Bjorvatn, A., Wörle, J., . . . Streubel, T. (2020). *L3 Pilot*

- Deliverable D3.3. Evaluation methods*. Retrieved from
- Neubauer, C., Matthews, G., & Saxby, D. (2014). *Fatigue in the automated vehicle: do games and conversation distract or energize the driver?* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.
- Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human factors*, 52(3), 381-410.
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human factors*, 39(2), 230-253.
- SAE. (2021). J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (Vol. J3016). www.sae.org: SAE International/ISO.
- Schömig, N., Hargutt, V., Neukum, A., Petermann-Stock, I., & Othersen, I. (2015). *The interaction between highly automated driving and the development of drowsiness*. Paper presented at the 6th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences, AHFE 2015, Las Vegas, USA.
- Seppelt, B., & Victor, T. (2020). Driver's mental model of vehicle automation. *Handbook of Human Factors for Automated, Connected, and Intelligent Vehicles*, 55-66.
- UNECE. (2021). *Addendum 156 - UN Regulation No. 157. Uniform provisions concerning the approval of vehicles with regard to Automated Lane Keeping Systems (E/ECE/TRANS/505/Rev.3/Add.156)*.
- Vogelpohl, T., Kühn, M., Hummel, T., & Vollrath, M. (2019). Asleep at the automated wheel—Sleepiness and fatigue during highly automated driving. *Accident Analysis & Prevention*(126), 70-84. doi:<https://doi.org/10.1016/j.aap.2018.03.013>

Assessment of the effectiveness of several countermeasures in reducing driver fatigue and associated risks for safety during autonomous driving.

Stephanie Dabic^{1*}, Min Luo¹; Manuel Dupont², Pierre Duysinx

¹ Valeo Switches & Smart Controls Product Line, Rue Jules VERNE - Vétraz Monthoux
BP 509 - 74106 Annemasse Cedex, FRANCE

*stephanie.dabic@valeo.com

² Second University of Liège, Allée de la Découverte 13A, Bât B52, 4000 Liège, BELGIQUE

Abstract: Driving an autonomous vehicle gradually induces Passive Fatigue (PF), and is likely to compromise safety when the driver is required to take back manual control. Literature suggests that thermal stimulation, auditory stimulation and secondary tasks engaging during manual driving are countermeasures that mitigate PF. However, no study to date attempted to replicate these results in autonomous driving conditions. To this end, we plan to conduct a randomized controlled experiment on a high-fidelity driving simulator (n=100). The study will employ a between-group design with the ‘type of countermeasure’ as an independent variable. To test the effect of those countermeasures on PF, five experimental groups will therefore be created, i.e. Control group; Thermal group with 4-min cooling at 18°C; Audio group with music listening group; Trivia group with secondary task, and Light condition with blue and red light. PF will be assessed using the Karolinska Sleepiness Scale and ocular parameters, while driving performance will be measured based on reaction times to a safety hazard. Finally, user comfort will be evaluated using the comfort Likert scale in order to determine which of the five countermeasures leads to the best compromise between risk-reduction and comfort. The aim of this study is to determine the effectiveness of the countermeasures described above in reducing driver fatigue and associated risks for safety while preserving satisfactory levels of comfort. 100 participants will be recruited. Results are in progress but in this paper we show first results on fatigue performance and driving performance. Electrocardiogram data analysis is in progress in order to find cardiac or respiratory signatures of PF before behavioral signs.

1. Introduction

The Take Over (TO) situations are critical and currently the focus of research. After a long period of autonomous driving, the driver must be in a state that allows him to take back control of the vehicle in a safe and secure manner. So it is crucial to look into the notion of Active Fatigue (AF) and Passive Fatigue (PF).

Indeed, several studies have also shown performance decrements in situations of cognitive overload or underload. Following these studies, Desmond and Hancock (2001) proposed a differentiated model of cognitive fatigue: active or passive. AF would be associated with cognitive overload whereas PF derived from underload and monotony. Workload reduction may be effective only for AF with PF calling for different countermeasures (May & Baldwin, 2009). According to this model, in driving situations AF results from “continuous and prolonged task-related perceptual-motor adjustments,” while PF results from, “system monitoring with either rare or no overt perceptual-motor response requirements” (Desmond & Hancock, 2001, p. 455).

In concrete terms, we could say that AF occurs when a person manually controls a system for extended periods of time, while PF occurs during periods of underload, requiring infrequent use of the controls (Eisert, 2018; Bernhardt et al., 2019; May & Baldwin, 2009; Matthews & Desmond, 2002). Given that driving automation can induce underload conditions, PF may be more of a concern than AF to operators

using automation in modern operational environments (Bernhardt et al., 2019; Saxby et al., 2013).

Driving an autonomous vehicle gradually induces PF, and is likely to compromise safety when the driver is required to take back manual control. Literature suggests that thermal stimulation, auditory stimulation and engaging in secondary tasks during manual driving are countermeasures that mitigate PF (Navarro and al, 2019). However, no study had attempted to replicate these results in autonomous driving conditions.

2. Method

2.1 Setup

To this end, we conducted a randomized controlled experiment on a high-fidelity driving simulator on 100 participants. The study employed a between-group design with the type of countermeasure as an independent variable (i.e. thermal stimulation, auditory stimulation, secondary task engaging, light countermeasure).

For thermal stimulation, the factor was tested at one level: 4 minutes cooling at 18° C, because 15 degrees is not comfortable for the user (Landstrom et al, 1999). The cooling duration and temperature are chosen based on previous studies (Schmidt, et al, 2017; Schmidt, Bullinger, 2017). The thermo-neutral climate was maintained (i.e. 24°C). To achieve cooling, we simulated the air conditioner to simulate a car vent that blows the wind towards the face, because facial stimulation is effective (Collins, and al 1996; Dalton, and al

2007; Health and als, 1990). Cooling began 4 minutes before the take-over request (see Figure 1).



Figure 1. Illustration of simulator setup

For the auditory stimulation, participants were asked to bring a playlist composed of heteroclite songs they like (Dalton, and al 2007). Music was played at comfort level immersively in the carriage (around 75 dBA) during the entire driving (Turner and als, 2017; Vogelpohl and als, 2019).

For light stimulation, an ingenious and innovative countermeasure, we had red and blue LED lights around the cabin, as discovered by Elloit 2015, who thinks light could be a promising way to reduce fatigue.

For the secondary task engaging condition, participants were asked to play the “TRIVIA game”, which consists of multiple choice questions (with 3 or 4 possible answers) that were read out and also displayed on a screen. Participants responded by pressing the buttons located on the center of the steering wheel.

To test the effects of those countermeasures on PF, five experimental groups therefore was created, i.e. control group (Control); 4-min cooling at 18°C (Thermal) ; music listening group (Audio); static red and blue light (Light); secondary task engaging group (Trivia).

PF was assessed using the Karolinska Sleepiness Scale (KSS) and ocular parameters, while driving performance was measured based on reaction times to a safety hazard. Finally, user comfort was evaluated using the Bedford scale in order to determine which of the four countermeasures leads to the best compromise between risk-reduction and comfort.

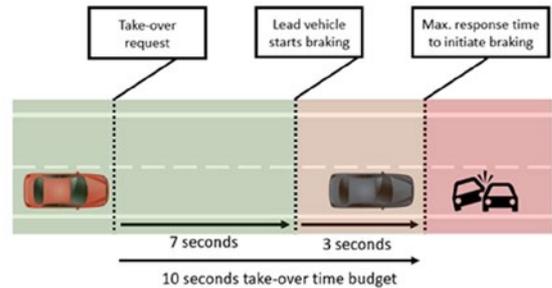


Figure 2. Illustration of the safety hazard

2.2 Purpose

The aim of this study is to determine the effectiveness of the countermeasures described above in reducing driver fatigue and associated risks for safety while preserving satisfactory levels of comfort. 100 participants were recruited. Take Over (TO) situation was described in Figure 2, occurring 40 min after the beginning without informing participants.

First, individuals willing to participate answered an online form consisting of the Pittsburgh Sleep Quality Index (PSQI) and a demographic questionnaire. Selected participants demonstrated good sleep quality (PSQI score < 6) and no regular intake of nicotine or drugs likely to interfere with cognitive functioning.

Then, participants were matched into blocks of twenty individuals according to demographic criteria (i.e. same gender, similar age and driving experience). Within each block, participants were randomly assigned to one of the five experimental conditions in order to control for these demographic criteria. All participants were paid 30€ upon completion of the experiment.

3. Results

In this study, ANOVA always used the type of countermeasure as a grouping factor, and only changed the dependent variable (DV).

3.1. KSS

ANOVA: when T0*KSS as a DV shows no significant difference ($F(4,95)=1.865, p>.001$). So, the KSS score has no statistical difference for the five groups before the beginning of the scenario.

ANOVA: when T35*KSS as a DV shows significant difference ($F(4,95)=2.607, p<.001, \eta^2p=0.099$). Post Hoc Tests reveal an almost significant difference between Audio and Control ($t=2,687 ; p<0.05$) show audio better for awakeness.

ANOVA: when ΔKSS (difference between KSS'0 and KSS'35) as a DV shows significant difference ($F(4,95)=4.357, p<.001, \eta^2p=0.155$). Post Hoc Tests revealed a significant difference between: Audio vs Light ($t=2,98; p<0.05$) Trivia vs Control ($t=2,775; p<0.05$) and Trivia vs

Light ($t=3,288$; $p<0.05$).

3.2. Pupils diameter

The objective measure of fatigue where we recorded 120'' before TO. The more the pupils are dilated, the more the participant is "awake".

ANOVA: when pupil diameter as DV shows no significant difference ($F(4, 95) = 1.656$, $p = 0.167$, $\eta^2p = 0.065$).

3.3. Fatigue questionnaire

The Fatigue questionnaire included 4 dimensions : Physical fatigue, mental fatigue, motivation and stress with 5-points Likert scales.

On Physical fatigue, a mix ANOVA when Time as repeated factor (before vs after) shows significant difference ($F(4,95) = 4.854$, $p < .001$, $\eta^2p = 0.018$), no significant effect of mental fatigue ($F(4,95) = 0.769$, $p > .05$) and Stress ($F(4, 95) = 0.064$, $p > .05$).

On Motivation, a mix ANOVA with Time as repeated factor (before vs after) shows significant difference ($F(4,95) = 33.531$, $p < .001$, $\eta^2p = 0.113$).

3.4. Take over performances

3.4.1. Time to take over the steering wheel after the TO Request

ANOVA: when time to reach the steering wheel as DV shows significant difference ($F(4,95) = 3.422$, $p < .001$, $\eta^2p = 0.126$). Post Hoc tests reveal significant differences between : Audio vs Trivia ($t = 2,957$; $p < 0.05$) Light vs Trivia ($t = 3,116$; $p < 0.05$).

3.4.2. Time to take over the pedals (accelerator or brake)

ANOVA: when time to reach the pedals as a DV shows no significant difference ($F(4,95) = 1.127$, $p = 0.349$, $\eta^2p = 0.045$).

3.4.3. First time watching the middle and side rear view mirrors after the TO request

ANOVA: when first time to watch the middle rear view mirror as a DV shows no significant difference ($F(4,50) = 0.615$, $p = 0.654$).

ANOVA: when first time to watch the left side rear view mirror as a DV shows significant difference : ($F(4,38) = 2.994$, $p = 0.030$) : Audio with longer Control and Light condition.

Many participants did not look at the middle mirror in the 60-second interval after TO. On average, participants watched it after the critical event.

3.4.4. Agreeability rating of the countermeasures

We do student tests on the Likert scale (7 points, 1 not pleasant, 7 very pleasant).

Audio and Trivia condition are significant the most pleasant ($F(4, 95) = 14.189$, $p < .001$, $\eta^2p = 0.359$), Thermic condition is neither pleasant nor unpleasant with all others conditions ($p > 0.05$) and Light is Very slightly unpleasant ($F(4, 95) = 37.93$, $p < .001$, $\eta^2p = 0.517$) (Figure 3).

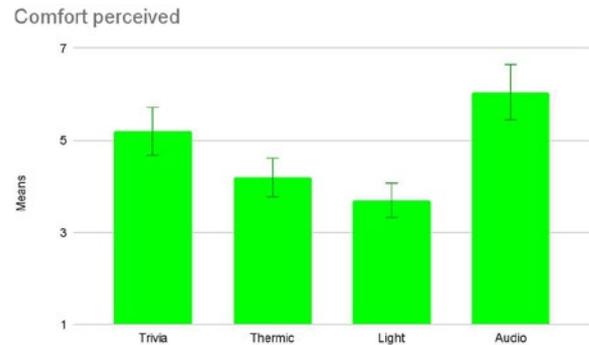


Figure 3. Comfort perceived in function condition of countermeasure.

4. Discussion

4.1. Impact of driving scenario and countermeasures on Fatigue

Increased KSS scores for each group. Significant impact difference on KSS scores. No significant impact on the objective measure of fatigue (pupils diameter). This study confirms that autonomous driving increases fatigue (Induction). Listening to music and playing games seem to be promising ways to counteract this fatigue.

4.2. Fatigue questionnaire

Regardless of the group, participants were slightly more physically tired and less motivated after the driving scenario than before, but not affected by remediation conditions.

4.3. Impact of countermeasures on Take Over performances

Participants in TRIVIA conditions were slower to take control back of the steering wheel and then think about the pedals later.

Watching the middle and side rear view mirrors is not a priority for drivers because they see them between 13 and 40 seconds after TO request, and more than half of the participants didn't look at it.

4.4. Impact of countermeasures on Reaction to the critical events

More collisions in the Control group show no

difference between the groups concerning the way to avoid the accident, but Control participants were slower to react to the critical event.

4.5. Agreeability of countermeasure

Audio and Trivia were judged more pleasant than other conditions significantly : they are promising ways to counteract this fatigue as previous results (Navarro et al 2019).

5. Conclusions and perspectives

On subjectives fatigue, Audio and Trivia are the best countermeasures. Indeed, few of the countermeasures : Audio longer for first glaze and rear mirror in less time to take over on steering wheel : perhaps overconfidence. We need to test more in detail in future new experiments.

Moreover, for reaction to the critical event : participants of the control group seem to have had more difficulties to properly react. So countermeasures have a positive effect on performance of TO.

Finally, Trivia conditions are the best compromise between risk-reduction and driver comfort. But it could involve an immersion problem.

This experiment is one of the first studies to examine the impact of fatigue due to a long period of autonomous driving on TO performances. Most of the studies : monotonous driving scenario but not autonomous. We plan to more control Take Over (TO) situations, situations more critical (eg less visual information to prevent driver, more difficult...).

Finally we need to investigate other promotive remediation and analyse ECG to find PF signature and predict it.

Acknowledgements

The analysis of this research was conducted with French financial support of BPI in the SERENITE project , without this study and communication would not have been possible.

References

Bernhardt, K. A., Poltavski, D., Petros, T., & Richard Ferraro, F. (2019, November). Differentiating Active and Passive Fatigue with the use of Electroencephalography. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 63, No. 1, pp. 1285-1289). Sage CA: Los Angeles, CA: SAGE Publications.

Collins, K. J., Abdel-Rahman, T. A., Easton, J. C., Sacco, P., Ison, J., & Dore, C. J. (1996). Effects of facial cooling on elderly and young subjects: interactions with breath-holding and lower body negative pressure. *Clinical Science*, 90(6), 485-492.

Dalton, B. H., & Behm, D. G. (2007). Effects of noise and music on human and task performance: A systematic review. *Occupational ergonomics*, 7(3), 143-152

Eisert, J. L. (2018). Using Theoretical Models of Fatigue to Design Optimal Fatigue Countermeasures (Doctoral dissertation, George Mason University).

Hancock, P. A., & Desmond, P. A. (2001). Stress, workload, and fatigue. Lawrence Erlbaum Associates Publishers.

Heath, M. E., & Downey, J. A. (1990). The cold face test (diving reflex) in clinical autonomic assessment: methodological considerations and repeatability of responses. *Clinical science*, 78(2), 139-147.

Landström, U., Englund, K., Nordstrom, B., & Stenudd, A. (1999). Laboratory studies on the effects of temperature variations on drowsiness. Perceptual and motor skills, 89(3_suppl), 1217-1229.

Navarro, J., Osiurak, F., Gaujoux, V., Ouimet, M. C., & Reynaud, E. (2019). Driving under the influence: how music listening affects driving behaviors. *JoVE (Journal of Visualized Experiments)*, (145), e58342.

May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transportation research part F: traffic psychology and behaviour*, 12(3), 218-224.

Matthews, G., & Desmond, P. A. (2002). Task-induced fatigue states and simulated driving performance. *The Quarterly Journal of Experimental Psychology: Section A*, 55(2), 659-686.

Saxby, D. J., Matthews, G., Warm, J. S., Hitchcock, E. M., & Neubauer, C. (2013). Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. *Journal of experimental psychology: applied*, 19(4), 287.

Schmidt, E., Decke, R., Rasshofer, R., & Bullinger, A. C. (2017). Psychophysiological responses to short-term cooling during a simulated monotonous driving task. *Applied ergonomics*, 62, 9-18.

Schmidt, E., & Bullinger, A. C. (2019). Mitigating passive fatigue during monotonous drives with thermal stimuli: Insights into the effect of different stimulation durations. *Accident Analysis & Prevention*, 126, 115-121.

Turner, M. L., Fernandez, J. E., & Nelson, K. (1996). The effect of music amplitude on the reaction to unexpected visual events. *The Journal of General Psychology*, 123(1), 51-62.

Vogelpohl, T., Kühn, M., Hummel, T., & Vollrath, M. (2019). Asleep at the automated wheel—Sleepiness and fatigue



during highly automated driving. Accident Analysis & Prevention, 126, 70-84.

Who benefits from napping in automated driving? – Effects of chronotype on subjective sleep inertia.

Markus Tomzig^{1*}, Christina Kaß¹, Johanna Wörle¹, Martin Baumann²

¹ Würzburger Institut für Verkehrswissenschaften (WIVW) GmbH, Robert-Bosch-Straße 4, Veitshöchheim, Germany

² Department of Human Factors, Ulm University, Ulm, Germany

*tomzig@wivw.de

Abstract: Sleeping during a trip is a promising feature of highly automated driving systems. However, sleep inertia may reduce the hedonic benefit of taking a nap. We investigated effects of time of day, individual chronotype, and driving time on subjective sleep inertia during highly automated driving in a simulator study. N = 20 participants (half morningness-, half eveningness-chronotypes) completed one driving session in the early morning and another in the late evening. Both sessions consisted of two manual drives, interrupted by a 50-minute nap while driving in automated mode. Participants indicated their subjective arousal, wellbeing and motivation to continue manual driving every five minutes. Results show not only that the participants felt worse in the evening than in the morning but also that arousal, wellbeing, and motivation were reduced if the chronotype did not fit the time of day. In the manual drive before the nap, arousal, wellbeing and motivation to drive decreased throughout the drive. After the nap, they remained on a low, but stable level. Napping during high automated driving thus provides a certain benefit by averting a further impairment – especially for morningness types who take a nap in the early morning.

1. Introduction

In contrast to partially or conditionally automated driving, high driving automation (SAE Level 4; SAE, 2021) will enable the driver to completely refrain from the driving task. Sleeping during the trip is one of multiple conceivable use cases of highly automated driving (Becker et al., 2018; Kyriakidis, Happee & de Winter, 2015) and promises recovery from sleepiness (Hartzler, 2014; Milner & Cote, 2009)

However, level 4 equipped vehicles might request the driver to take over the driving task at the end of the Operational Design Domain (ODD; SAE, 2021). Estimated effects of sleep recovery are likely to be opposed by the phenomenon of sleep inertia which is defined as “grogginess, disorientation, and sleepiness that can accompany awakening from a nap” (Rosekind et al., 1995). Former studies demonstrated that sleep inertia impairs performance in various tasks (Ferrara & De Gennaro, 2000; Hilditch & McHill, 2019; Tassi & Muzet, 2000). In the context of automated driving, take over performance is worsened by sleep inertia (Wörle, Metz & Baumann, 2021; Wörle, Metz, Othersen & Baumann, 2020). The extent of sleep inertia is influenced by several external factors, like the time of day (Hilditch & McHill, 2019; Tassi & Muzet, 2000), as well as individual factors, like the personal chronotype (Ma et al., 2022).

Previous research focused mainly on investigating the effects of sleep inertia on task performance and driving behaviour. However, it seems plausible that sleep inertia affects also the subjective wellbeing negatively and thus reduces the hedonic benefit of a nap. This in turn might prevent drivers taking their naps even if possible and thereby increase the risk of fatigue induced accidents when the vehicle leaves the level 4 ODD. Therefore, the aim of the presented study was to assess subjective sleep inertia

during highly automated driving in dependence of the time of day and chronotype.

2. Method

2.1 Procedure

The participants were invited for three experimental driving sessions in the high-fidelity driving simulator of the Würzburg Institute for Traffic Sciences (WIVW). To assess influences of the time of day in the context of sleep inertia, one of the three drives took place in the early morning (6 a.m.), another in the late evening (9 p.m.). The order of the three sessions was partially randomized. Procedure and results of the third drive are reported elsewhere.

To ensure that participants were sleepy at the beginning of the sessions, they were allowed to sleep no more than four hours in the night before taking part in each of the experimental drives. The drives started with a 30 minute manual drive on a motorway with a (fictional) speed limit of 180 km/h. After that, participants were instructed to activate the highly automated driving system and to sleep. After 50 minutes of automated driving, participants were awakened by a request to intervene, asking the participants to take over the driving task promptly. The takeover was followed by a second 30 minute manual drive which was identical to the first manual drive.

2.2 Measures

At the beginning and every five minutes during the manual drives, participants were asked to indicate their subjective arousal (“How activated do you feel at the moment?”), wellbeing (“How well do you feel at the moment?”), and motivation to continue the manual drive (“To what degree would you like to continue the manual drive at the moment?”). The participants answered the

questions verbally on 9-point Self Assessment Manikin (SAM) scales (Bradley & Lang, 1994; depiction retrieved from Laghari et al., 2013) and in total seven times before and seven times after sleeping. As an exception, the first inquiry after the takeover was not conducted immediately but retrospectively, five minutes afterwards in order not to bias the driving behaviour during takeover. The wording of the questions was adapted respectively (“How activated did you feel five minutes ago immediately after takeover?” and analogously for wellbeing and motivation).

During the manual drives, different measures of driving and glance behaviour, heartrate, and performance in an auditory vigilance task were collected but are reported elsewhere. Sleep was measured with electroencephalography (EEG; Brain Products GmbH) and confirmed according to the American Academy of Sleep Medicine standard (AASM, 2017).

2.3 Sample

Previous to the study, $N = 183$ participants were screened with the German version of the reduced Morningness-Eveningness Questionnaire (rMEQ; Randler, 2013). $N = 20$ participants (7 female, mean age = 42, $SD = 15$) were selected according to their individual chronotype and took part in the study ($n = 10$ were morningness types with rMEQ score ≥ 18 and $n = 10$ were eveningness types with rMEQ score ≤ 11). All participants held a valid driving license and had completed an extensive driving simulator training.

2.4 Data Analysis

Univariate mixed Analyses of Variance (ANOVA) for each of the three subjective measures were performed to examine the impact of different factors on arousal, wellbeing and motivation to continue the drive in the context of sleepiness and sleep inertia. Time of day (2-level within-subjects factor), chronotype (2-level between-subjects factor), driver state (sleepiness vs. sleep inertia; 2-level within-subjects factor), and number of inquiry during the manual drive (7-level within-subjects factor) were included as independent variables in the model.

3. Results

19 of 20 participants experienced EEG-verified sleep in the morning drive, 18 of 20 in the evening drive. The univariate ANOVAs showed that the participants felt significantly worse in the evening (lower wellbeing) and were less motivated to drive manually than in the morning. There was no statistically significant difference concerning the mean arousal.

There was no significant main effect of chronotype but significant interaction effects between the time of day and the chronotype for all three dependent variables. As can be seen in figure 1, arousal differed mainly in the morning, wellbeing in the evening, and motivation for both times of day between the two chronotypes. For all three measures, this effect derived mainly from the morning types, differing between the times of day, whereas the evening types did sparsely distinguish between morning and evening drives. Arousal, wellbeing and motivation were thus better if the time of day did fit to the chronotype for morningness types, but not for eveningness types.

Across all conditions, the arousal, wellbeing, and motivation decreased before sleeping but remained more or less stable after sleeping. Accordingly, the interaction effects between driver state and number of inquiry were significant for arousal and motivation and marginally significant for wellbeing. On a descriptive level, only in the morning drive, the arousal of the morningness types improved slightly after sleeping. All statistical results are listed in table 1.

4. Discussion

Sleep is a desired use case of high driving automation. In our study, participants were instructed to sleep during two simulator drives and were requested to drive manually before and after sleeping. We examined the time course of the subjective arousal, wellbeing and motivation to drive during manual driving in dependence of time of day, driver state, and individual chronotype.

Overall, there was no benefit of sleeping during high automated driving in the proper sense as arousal, wellbeing and motivation to continue the manual drive did not improve after sleeping compared to before sleeping. On closer look, however, sleep seemed to avert a further decrease of the subjective ratings. Further, the ratings differed in

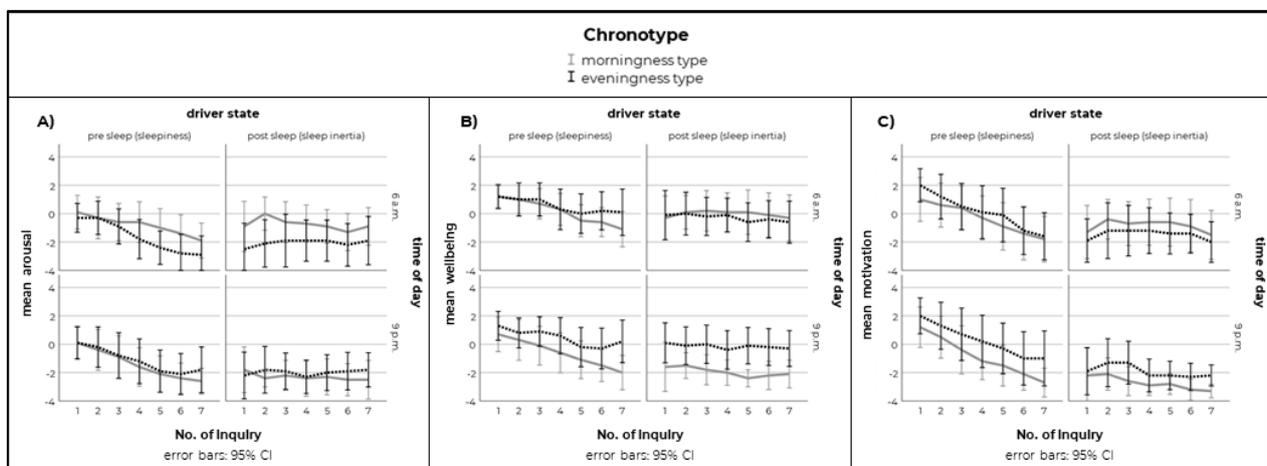


Fig. 1: Time courses of A) arousal, B) wellbeing, and C) motivation across the experimental conditions.

Table 1: Results of the univariate mixed ANOVAs.

| Effect | Measure | <i>F</i> | <i>p</i> | η_p^2 |
|---|------------|--------------------|----------|------------|
| Main effect: time of day | Arousal | $F(1, 18) = 2.727$ | .116 | .132 |
| | Wellbeing | $F(1, 18) = 4.564$ | .047 | .202 |
| | Motivation | $F(1, 18) = 7.405$ | .014 | .291 |
| Main effect: chronotype | Arousal | $F(1, 18) = 0.326$ | .575 | .018 |
| | Wellbeing | $F(1, 18) = 1.546$ | .230 | .079 |
| | Motivation | $F(1, 18) = 0.374$ | .549 | .020 |
| Interaction effect: time of day * chronotype | Arousal | $F(1, 18) = 8.135$ | .011 | .311 |
| | Wellbeing | $F(1, 18) = 5.079$ | .037 | .220 |
| | Motivation | $F(1, 18) = 4.493$ | .048 | .200 |
| Interaction effect: driver state * no. of inquiry | Arousal | $F(6, 13) = 4.643$ | .010 | .682 |
| | Wellbeing | $F(6, 13) = 2.864$ | .053 | .570 |
| | Motivation | $F(6, 13) = 5.420$ | .005 | .714 |

dependence of the individual chronotype and time of day: Whereas eveningness types rated their arousal, wellbeing and motivation on a comparable level in the morning and in the evening, morningness types distinguished between both times of day. In the morning drive, the arousal of the morningness types even improved slightly after sleeping.

Limiting factors like the impact of partial sleep deprivation as methodological tool to induce sleepiness must be discussed.

Further analyses are planned to assess whether the subjective wellbeing after sleep depends on the depth of sleep which is measured by EEG. Additional analyses will reveal if and how sleepiness and sleep inertia affect driving behaviour and to what extent this correlates with the participants' subjective wellbeing.

5. Conclusions

In our study, sleeping during highly automated driving did not improve the subjective arousal, wellbeing and motivation in the subsequent manual drives. Nonetheless, sleeping provides a certain benefit by averting a further impairment – especially if morningness types took a nap in the early morning.

6. Acknowledgments

This study was conducted within the project RUMBA (for more information see <https://projekt-rumba.de/>) which receives funding from the German Federal Ministry for Economic Affairs and Climate Action. The sole responsibility of this publication lies within the authors. The authors would like to thank all partners within RUMBA for their cooperation and valuable contribution.

References

AASM. (2017). *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications.*: American Academy of Sleep Medicine.

Becker, T., Herrmann, F., Duwe, D., Stegmüller, S., Röckle, F., & Unger, N. (2018). Enabling the value of time.

Implications for the interior design of autonomous vehicles: Fraunhofer IAO.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1), 49-59.

Ferrara, M., & De Gennaro, L. (2000). The sleep inertia phenomenon during the sleep-wake transition: theoretical and operational issues. *Aviation, space, and environmental medicine*, 71(8), 843-848.

Hartzler, B. M. (2014). Fatigue on the flight deck: the consequences of sleep loss and the benefits of napping. *Accident Analysis & Prevention*, 62, 309-318.

Hilditch, C. J., & McHill, A. W. (2019). Sleep inertia: current insights. *Nature and science of sleep*, 11, 155.

Kyriakidis, M., Happee, R., & de Winter, J. C. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation research part F: traffic psychology and behaviour*, 32, 127-140.

Laghari, K., Gupta, R., Falk, T., Antons, J., Schleicher, R., & Möller, S. (2013). *Objective characterization of human behavioural characteristics for QoE assessment: A pilot study on the use of electroencephalography features.* Paper presented at the 2013 IEEE Globecom Workshops, GC Wkshps 2013.

Ma, Z., Chen, X. Y., Wang, D., Zhu, Z., Niu, H., Huang, S., . . . Fan, F. (2022). Who is the hardest to wake up from sleep? An investigation of self-reported sleep inertia using a latent profile analysis. *Journal of sleep research*, e13552.

Milner, C. E., & Cote, K. A. (2009). Benefits of napping in healthy adults: impact of nap length, time of day, age, and experience with napping. *Journal of sleep research*, 18(2), 272-281.

Randler, C. (2013). German version of the reduced Morningness-Eveningness Questionnaire (rMEQ). *Biological Rhythm Research*, 44(5), 730-736.

- Rosekind, M. R., Smith, R. M., Miller, D. L., Co, E. L., Gregory, K. B., Webbon, L. L., . . . Lebacqz, J. V. (1995). Alertness management: strategic naps in operational settings. *Journal of sleep research, 4*, 62-66.
- SAE. (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles* (Vol. J3016): SAE International.
- Tassi, P., & Muzet, A. (2000). Sleep inertia. *Sleep medicine reviews, 4*(4), 341-353.
- Wörle, J., Metz, B., & Baumann, M. (2021). Sleep inertia in automated driving: Post-sleep take-over and driving performance. *Accident Analysis & Prevention, 150*, 105918.
- Wörle, J., Metz, B., Othersen, I., & Baumann, M. (2020). Sleep in highly automated driving: Takeover performance after waking up. *Accident Analysis & Prevention, 144*, 105617.

Heart rate variability as an indicator for driver fatigue, different effects of time of day and time-on-task

Ke Lu ^{1,2*}, Anna Sjörs Dahlman ^{1,2,3}, Johan Karlsson ^{2,4}, Bengt Arne Sjöqvist ^{1,2}, Stefan Candefjord ^{1,2}

¹ Department of Electrical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden

² SAFER Vehicle and Traffic Safety Centre, Chalmers University of Technology, 41756 Gothenburg, Sweden

³ Swedish National Road and Transport Research Institute (VTI), 58330 Linköping, Sweden

⁴ Autoliv Research, Autoliv Development AB, 44737 Vårgårda, Sweden

*ke.lu@chalmers.se

Abstract: Heart rate variability (HRV) has been considered as a potential physiological marker for driver fatigue. However, consensus has not been reached for how HRV changes during the development of fatigue, due to inconsistent results in the literature. One potential cause for inconsistent results is that different causal factors were used to introduce fatigue. The aim of this study is to investigate how HRV parameters change during driving in relation to fatigue caused by sleep related and task related factors. Data from a real road experiment, with 89 participants who drove four times over a 180 km route, were used for the analysis. We investigated how time of day and time-on-task factors influence HRV parameters. The result shows that different HRV parameters react differently in relation to time of day and time-on-task factors. The result emphasizes the importance of considering the causal factors when interpreting results from driver fatigue studies and when developing fatigue detectors based on physiological measures.

1. Introduction

Physiological measurements have potential to complement conventional driver monitoring systems based on driving performance and facial features when facing the challenges of application of automated driving systems (Gonçalves & Bengler, 2015). Heart rate variability (HRV), a physiological marker reflecting changes of the cardiac sympathetic and parasympathetic branches of the autonomic nervous system, is of interest for detecting drivers' fatigue during real life driving (Lohani et al., 2019).

Many studies have investigated the relation between driver fatigue and HRV (Buendia et al., 2019; Jung et al., 2014; Lenis et al., 2016; Li & Chung, 2013; Lu et al., 2021; Patel et al., 2011; Persson et al., 2020; Vicente et al., 2016). However, inconsistent results can be found regarding how HRV parameters change when the driver transitions from alert to fatigued state. Several studies suggested LF/HF to be an important indicator of fatigue, as a reflection of the balance between parasympathetic and sympathetic nerve activity. Both increased and decreased LF/HF, as well as changes in several other HRV features, have been reported when drivers become fatigued. One potential cause of the inconsistency can be that different interventions were used to cause fatigue.

Fatigue is a complex phenomenon caused by multiple factors. It has been suggested that driver fatigue has both sleep related and task related causes (May & Baldwin, 2009). Sleep related fatigue is influenced by the circadian rhythm and the sleep homeostat, which depends on sleep duration and time awake since last sleep episode. Task related fatigue depends on time-on-task as well as the mental task load, where both underload and overload can contribute to fatigue.

Studies have taken different approaches to introduce fatigue to drivers. For sleep related fatigue, some studies used different time of day to perform driving tasks, and some have introduced partial or complete sleep deprivation before the driving session. When it comes to task related fatigue, some

studies have opted to use monotonous driving tasks to speed up the development of fatigue, and some studies have relied on the time-on-task effect for development of fatigue with continuous and prolonged driving tasks.

The purpose of this study is to investigate relationships between HRV parameters and sleep related and task related factors. In this study we used data from real road motorway driving in both afternoon and night to investigate the HRV changes associated to time of day and time-on-task factors.

2. Method

2.1 Dataset

The dataset consists of driving data from 89 drivers (36 female and 53 male). Each participant had four driving sessions in two different days. In each day, the participant had one driving session in the afternoon (starting at 15:00 or 17:00) and one at night (starting at 01:00 or 03:00). On one day the driver drove with partial automated driving mode, and the other day with manual driving. The test route comprised a 90-km section of a dual-lane motorway (road E4, Sweden) where the participants drove from exit 111 to exit 104 and back, resulting in a 180 km drive. The posted speed limit was 120 km/h on the whole section.

Lead II electrocardiogram among several other physiological measurements was recorded with bio-amplifier with a sampling frequency of 512 Hz. The recordings were then down sampled and stored with 256 Hz.

Detailed description of the experiment and analysis with other physiological measures, partial automated driving condition and subjective sleepiness rating can be found in (Ahlström et al., 2021; Lu et al., 2021).

2.2 Signal Processing

PhysioNet cardiovascular signal toolbox was used for the ECG and HRV analysis (Vest et al., 2018). The RR

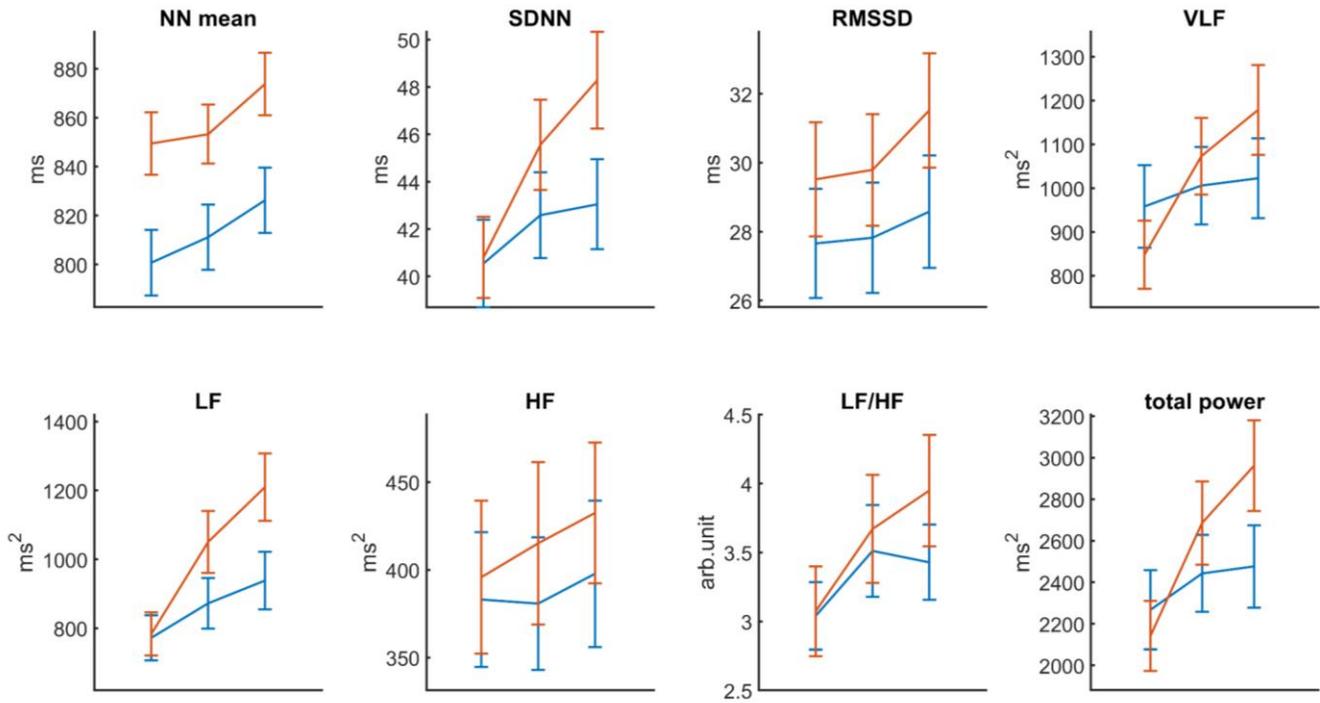


Fig. 1. Mean value of HRV features in three different time-on-task segments (x axis) and separated by afternoon (blue) and night (red) driving. Error bars represent standard error of mean.

intervals were extracted from the ECG measurement with visual inspection performed to remove cases with low signal quality that leads to wrong peak detection. HRV features were then extracted with 5-min wide sliding window with 1-min step size. Time domain features including NN mean, SDNN, RMSSD, and frequency domain features including VLF, LF, HF, LF/HF, and total power were extracted. The description of the features can be found in (Shaffer & Ginsberg, 2017). All data processing and analysis were performed with Matlab 2021b (MathWorks Inc., MA, USA).

2.3 Statistical Analysis

All 5-min epochs were grouped by time-on-task and day/night driving. For the time-on-task factor, each drive was divided into four 20-min long segments, two with outward and two with return driving. The first 20 min segment was excluded for the analysis to eliminate the influences at the start of the driving caused by being attached with sensors, talking to the test leader and getting familiar with the task, etc.

The influence of time of day and time-on-task on HRV metrics were analysed with two-way ANOVA test. The HRV parameters with skewed distribution were logarithmic transformed. For comparison between groups, paired T tests were performed. For these tests, the level of significance was set at $p < 0.05$, and Bonferroni correction for multiple testing was applied.

3. Results

The mean value and standard error of mean of HRV features in different time-on-task and time of day segments are shown in Fig 1.

The result of the two-way ANOVA test is shown in Table 1. We can find significant effect from time of day and time-on-task on almost all selected HRV features, except that the effect from time of day was not significant on VLF. In addition, the interaction of the two variables had significant effect on SDNN, VLF, LF, LF/HF and total power.

When it comes to comparisons between afternoon drive and the night drive for each time-on-task segment, most

Table 1 Results of ANOVA test.

| HRV features | Time of day (df=1) | | Time-on-task (df=2) | | Interaction, time of day * time-on-task | |
|--------------|--------------------|-------------------|---------------------|-------------------|---|-------------------|
| | F | p | F | p | F | p |
| NN mean | 298.96 | <0.0001 | 29.39 | <0.0001 | 0.46 | 0.6300 |
| SDNN | 40.10 | <0.0001 | 46.69 | <0.0001 | 7.19 | 0.0008 |
| RMSSD | 48.13 | <0.0001 | 7.65 | 0.0005 | 0.44 | 0.6443 |
| VLF | 2.36 | 0.1249 | 47.70 | <0.0001 | 10.37 | <0.0001 |
| LF | 41.15 | <0.0001 | 66.11 | <0.0001 | 8.02 | 0.0003 |
| HF | 11.24 | 0.0008 | 4.32 | 0.0134 | 0.50 | 0.6069 |
| LF/HF | 9.89 | 0.0017 | 59.88 | <0.0001 | 6.83 | 0.0011 |
| Total power | 14.87 | 0.0001 | 46.41 | <0.0001 | 8.79 | 0.0002 |

Bold field indicates $p < 0.0167$

features show significant differences, but there was no significant difference in mean values between afternoon and night for the first time-on-task segment for SDNN, LF, HF, and LF/HF. When comparing the first and last time-on-task segments in afternoon or night driving, significant increase can be found for all features except for HF in afternoon driving.

4. Discussion and Conclusions

In this study we investigated the influence of two fatigue factors on HRV parameters with data from a real road driving experiment. Different response patterns were found for different HRV parameters. For NN mean (heart rate) and RMSSD, time of day has a strong effect showing from the beginning of the driving, while the time-on-task has a smaller effect, and no significant interaction can be found between the two factors. For SDNN, LF and LF/HF, no significant difference can be found for afternoon and night driving at the beginning of the driving, but the night driving is causing steeper change on time-on-task compared to afternoon driving.

Fatigue caused by different factors has different influence on driver performance and countermeasures may vary accordingly (Williamson et al., 2011). The result from this study shows different HRV parameters react differently in relation to time of day and time-on-task factors. This difference may indicate a physiological difference in sleep related and task related fatigue. This knowledge will be valuable when interpreting results from related studies where different fatigue manipulation methods were applied. It also emphasizes the importance of considering the fatigue causing factors when designing future studies and developing fatigue detectors based on physiological measures.

5. Acknowledgments

This work was supported in part by the European Union's Horizon 2020 Research and Innovation Programme Projects Mediator under Grant 814735, in part by ADAS&ME under Grant 688900, in part by the Strategic Vehicle Research and Innovation (FFI) Program at the Sweden's Innovation Agency (VINNOVA) under Grant 2020-05157, and in part by Autoliv Development AB and SmartEye AB.

References

- Ahlström, C., Zemblys, R., Jansson, H., Forsberg, C., Karlsson, J., & Anund, A. (2021). Effects of partially automated driving on the development of driver sleepiness. *Accident Analysis & Prevention*, *153*, 106058.
- Buendia, R., Forcolin, F., Karlsson, J., Arne Sjöqvist, B., Anund, A., & Candefjord, S. (2019). Deriving heart rate variability indices from cardiac monitoring—An indicator of driver sleepiness. *Traffic Injury Prevention*, *20*(3), 249–254. <https://doi.org/10.1080/15389588.2018.1548766>
- Gonçalves, J., & Bengler, K. (2015). Driver State Monitoring Systems—Transferable Knowledge Manual Driving to HAD. *Procedia Manufacturing*, *3*, 3011–3016. <https://doi.org/10.1016/j.promfg.2015.07.845>
- Jung, S. J., Shin, H. S., & Chung, W. Y. (2014). Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel. *IET Intelligent Transport Systems*, *8*(1), 43–50. <https://doi.org/10.1049/iet-its.2012.0032>
- Lenis, G., Reichensperger, P., Sommer, D., Heinze, C., Golz, M., & Dössel, O. (2016). Detection of microsleep events in a car driving simulation study using electrocardiographic features. *Current Directions in Biomedical Engineering*, *2*(1), 283–287.
- Li, G., & Chung, W. Y. (2013). Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier. *Sensors (Switzerland)*, *13*(12), 16494–16511. <https://doi.org/10.3390/s131216494>
- Lohani, M., Payne, B. R., & Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. In *Frontiers in Human Neuroscience* (Vol. 13). <https://doi.org/10.3389/fnhum.2019.00057>
- Lu, K., Karlsson, J., Dahlman, A. S., Sjöqvist, B. A., & Candefjord, S. (2021). Detecting Driver Sleepiness Using Consumer Wearable Devices in Manual and Partial Automated Real-Road Driving. *IEEE Transactions on Intelligent Transportation Systems*, 1–10. <https://doi.org/10.1109/TITS.2021.3127944>
- May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(3), 218–224. <https://doi.org/10.1016/j.trf.2008.11.005>
- Patel, M., Lal, S. K. L., Kavanagh, D., & Rossiter, P. (2011). Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert Systems with Applications*, *38*(6), 7235–7242. <https://doi.org/10.1016/j.eswa.2010.12.028>
- Persson, A., Jonasson, H., Fredriksson, I., Wiklund, U., & Ahlstrom, C. (2020). Heart Rate Variability for Classification of Alert Versus Sleep Deprived Drivers in Real Road Driving Conditions. *IEEE Transactions on Intelligent Transportation Systems*, 1–10. <https://doi.org/10.1109/tits.2020.2981941>
- Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, *5*. <https://doi.org/10.3389/fpubh.2017.00258>
- Vest, A. N., Da Poian, G., Li, Q., Liu, C., Nemati, S., Shah, A. J., & Clifford, G. D. (2018). An open source benchmarked toolbox for cardiovascular waveform and interval analysis. *Physiological Measurement*, *39*(10), 105004.
- Vicente, J., Laguna, P., Bartra, A., & Bailón, R. (2016). Drowsiness detection using heart rate variability. *Medical and Biological Engineering and Computing*, *54*(6), 927–937. <https://doi.org/10.1007/s11517-015-1448-7>
- Williamson, A., Lombardi, D. A., Folkard, S., Stutts, J., Courtney, T. K., & Connor, J. L. (2011). The link between fatigue and safety. *Accident Analysis and Prevention*, *43*(2), 498–515. <https://doi.org/10.1016/j.aap.2009.11.011>

“I’m Sad When You’re Distracted” – Effectiveness and User Experience of an Innovative Driver Monitoring System for Partially Automated Driving

Christina Kaß^{1*}, Nadja Schömig¹, Yannick Forster², Frederik Naujoks², Katharina Wiedemann¹, Christian Maag¹, Andreas Keinath², Alexandra Neukum¹

¹ Würzburg Institute for Traffic Sciences (WIVW GmbH), Robert-Bosch-Str.4, 97209 Veitshöchheim, Germany

² BMW Group, Knorrstraße 147, 80937 München, Germany

*kass@wivw.de

Abstract: The presented simulator study investigated the effectiveness, user experience and usability of an innovative driver monitoring system (DMS) for partially automated driving, called “Jeannie”. This virtual assistant provided continuous visual emotional feedback dependent on drivers’ monitoring behaviour and issued warnings and speech outputs in response to prolonging eyes-off or hands-off driving. Furthermore, it supported the drivers maintaining an adequate attentional state, e.g. by relieving them from distracting tasks. Jeannie was compared to a more conventional DMS that only displayed warnings in the instrument cluster. The two DMS variants were comparably effective in preventing hands-off and eyes-off driving and where both highly accepted. However, user experience was higher for Jeannie.

1. Introduction

When driving with a partially automated system (Level 2 (L2); SAE, 2018) the driver is still responsible to monitor the roadway and to be ready to react to a system limit or error at any time. Therefore, driver monitoring systems (DMS) should assure that drivers keep their hands on the steering wheel and their eyes on the road, and warn if drivers do not fulfill their responsibilities. For the hands-on requirement, clear regulations exist how and when systems should intervene by issuing warnings (UNECE regulation R79, 2017). There are currently no such regulations for visual attention warnings. Naturalistic driving studies provided evidence that drivers may not adequately meet their responsibilities when using L2 systems. With active Tesla Autopilot (with hands-on requirement), drivers have taken their hands off the wheel more frequently than during manual driving (Morando et al., 2020). Blanco et al. (2015) found that visual inattention warnings encouraged drivers to monitor the road. However, over the course of the simulator study, some drivers became habituated and ignored the warnings to complete a non-driving related task.

A DMS should support drivers’ understanding of the responsibility for a continuous monitoring, should be accepted, and not perceived as paternalism. To identify the characteristics of a comprehensive and user-friendly DMS, we developed two different DMS. Both variants issued warnings in response to prolonging hands-off or eyes-off behavior. While the conventional DMS only issued discrete warnings in the cluster display, the innovative DMS additionally provided continuous visual emotional feedback to the driver. Dependent on drivers’ behavior, the emoji-like avatar called “Jeannie” continuously changed its emotions.

2. Method

2.1 Automated system and HMI

A prototypical L2 system was implemented in the static WIVW driving simulator. It included two distinct

modes: hands-off and hands-on driving. The system state was indicated in the instrument cluster display (Fig.1). It constantly displayed a symbol for L2 driving (green icon) and two separate indicators for drivers’ responsibilities: a steering wheel with either hands on or hands off the wheel (dependent on system mode) and an iconic eye symbol indicating the responsibility to monitor the roadway. An additional text box included the name of the system and displayed the drivers’ responsibilities (“driving assistant active; monitor traffic environment; keep hands on the wheel”).



Fig. 1. HMI for the L2 system in the instrument cluster.

2.2 Driver monitoring systems

Two different driver monitoring systems were implemented as a between subject factor: a conventional DMS and an innovative DMS. Both DMS used the same sensors and measured whether the drivers had their hands on the wheel during L2 hands-on mode and whether they kept their eyes on the road. Hands-on detection was implemented in the steering wheel. Eyes-off detection was measured with the SmartEye® system. Glances towards the instrument cluster, other areas inside the vehicle or too long intervals of tracking losses were defined as distraction.

The warning strategy of the conventional DMS for hands-off driving was based on the requirements of the UNECE R79: a yellow hands-off symbol was displayed as a visual warning if hands-off time exceeded 15s. After 30 seconds, a visual warning in red together with a speech output

(“please take hands on the wheel”) was triggered. Finally, after 60 seconds, if the driver has not responded to the previous warnings, a safe stop has been initiated. The warning strategy for eyes-off driving in the conventional DMS was defined as follows: if the driver was distracted for 4 seconds, the first warning stage was triggered (yellow eye icon), the second stage was triggered after 7s (red eye icon and speech output: “Monitor traffic situation”). A safe stop has been initiated after 10s.

The innovative DMS also used these warning stages and the respective HMI feedback in the instrument cluster but added a virtual assistant called “Jeannie”. Jeannie was shown on an additional display positioned in the upper part of the centre console. Besides the discrete warnings, it provided continuous feedback on drivers’ current monitoring and hands-off/on behavior by displaying various emotional states. If the drivers adequately monitored the roadway and had their hands on the wheel, Jeannie looked happy. As soon as the drivers took their attention off the road or took their hands off the wheel, Jeannie’s facial expression became neutral and subsequently turned to an unhappy state in several stages. If the drivers looked back to the road or grabbed the wheel, Jeannie first got neutral and after 5s happy again (Fig. 2). In contrast to the conventional DMS, Jeannie provided a more human-like speech output at the second warning stage. Another characteristic of Jeannie was that it provided drivers with specific support in dedicated use cases by offering voice

interaction (see chapter 2.4). Glances to Jeannie were not defined as distraction by the DMS.

2.3 Test sample

The sample consisted of 30 subjects (14 female). Mean age of the subjects was 41 years ($SD = 14.9$ years).

2.4 Test drives and test procedure

After very reduced instructions about the L2 system without information about the DMS, each driver performed four drives. In drive A, the drivers experienced L2 driving intuitively in several use cases for about 20 minutes (mainly driving in hands-off and hands-on mode). In drive B, specific use cases were introduced where drivers were explicitly instructed, e.g., to text with a friend via the smartphone. While drivers with the conventional DMS had to type the conversation, drivers with the innovative DMS were supported by Jeannie who provided voice interaction. In drive C, the drivers were explicitly instructed to direct their attention away from the road and to take their hands off the steering wheel in order to experience the DMS warning stages subsequently as well as the emotional feedback by Jeannie. In drive D, drivers again experienced the DMS warning stages but now via the DMS system from the other condition.

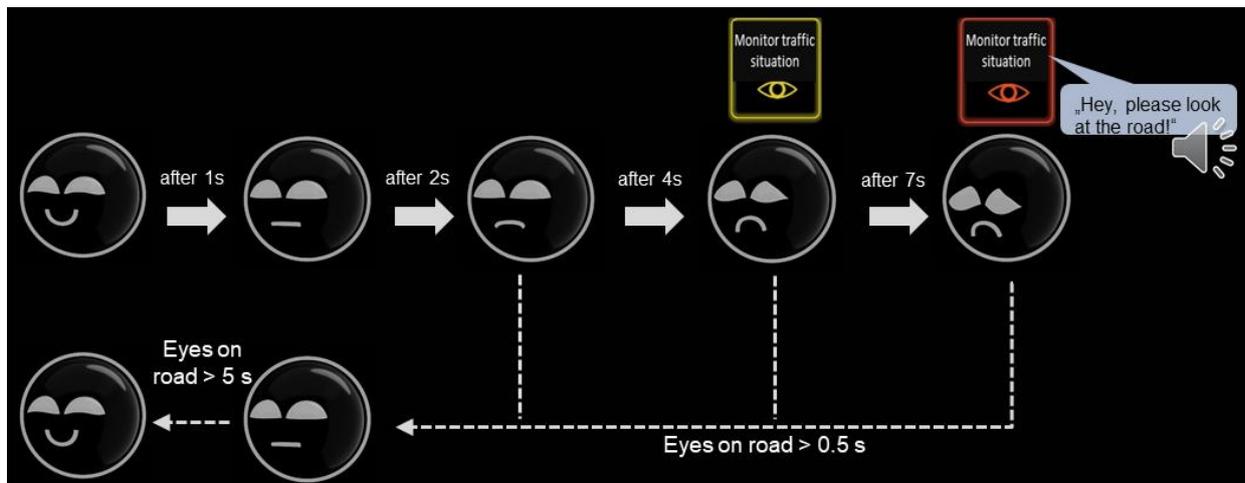


Fig. 2. Emotional stages of Jeannie based on visual distraction.

2.5 Dependent measures

As a measure for effectiveness, the number of received DMS eyes-off and hands-off warnings in drive A was counted. System usability was assessed via SUS (Brooke, 1996) after drive A and B. User experience was assessed via UEQ after drive A and drive B (Laugwitz et al., 2008). Finally, drivers should indicate whether or not they want Jeannie as a DMS for L2 driving.

3. Results

The comparison of the number of hands-off and eyes-off warnings showed that both DMS variants were equally effective to assure drivers’ adequate responsibility in L2 driving. Drivers received a comparable low number of hands-off warnings ($\chi^2 = 3.167, p = .367$). Eyes-off warnings were

generally more frequent as drivers explored the system and the HMI intensively during drive A. However, the number of warnings did not differ between the DMS variants ($\chi^2 = 11.667, p = .473$).

System usability measured by SUS was also comparably high in both conditions and reached very good to excellent values ($t(28) = -.243; p = .810$). In the Jeannie condition, usability increased on a descriptive level from drive A to B after experiencing the support functions from Jeannie. A multivariate ANOVA with the UEQ scales attractiveness, pragmatic, and hedonic quality as dependent variables revealed a significantly higher user experience of Jeannie (Wilks’ $\lambda = .716, F(3,26) = 3.44, p = .031, \eta_p^2 = .28$): Analyses of the subscales showed that this difference was mainly based on a higher hedonic quality of Jeannie which even increased from drive A to B (factor DMS: $F(1,28) = 6.84,$

$p = .014$, $\eta_p^2 = .20$; factor drive: $F(1,28) = 4.25$, $p = .049$, $\eta_p^2 = .13$).

When asked whether drivers would choose Jeannie or not in a partially automated ride, eight drivers in the Jeannie condition and six drivers in the conventional DMS would choose Jeannie. Drivers from the Jeannie condition preferred the supportive function but not the continuous emotional feedback, which some drivers found additionally distracting.

4. Discussion and conclusions

The results revealed that both DMS variants were comparably effective in preventing hands-off and eyes-off driving. Both variants were highly usable. Jeannie achieved a higher user experience. This rating was mainly based on the additional supportive functions and the more human-like speech interaction while the continuous emotional feedback was not that appreciated. The results can be used as motivation to create more innovative but still effective and accepted DMS solutions for L2 driving.

5. Acknowledgments

This research has been funded by the BMW AG. Statements and opinions expressed in this paper reflect that of the authors and not the BMW AG. Thanks to the reviewers for their valuable time and effort.

References

- Society of Automotive Engineers International J3016 (2018). Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Warrendale, PA: SAE International.
- UNECE, Addendum 78 (2017). UN Regulation No. 79, Revision 3, E/ECE/TRANS/505/Rev.1/Add.78/Rev.3, Geneva, Switzerland, 30 November 2017.
- Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J., ... & Morgan, J. F. (2015). Human factors evaluation of level 2 and level 3 automated driving concepts (Report No. DOT HS 812 182). Washington, DC: National Highway Traffic Safety Administration.
- Morando, A., Gershon, P., Mehler, B., & Reimer, B. (2020). Driver-initiated Tesla Autopilot disengagements in naturalistic driving. Proceedings of the 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 57-65. doi:10.1145/3409120.3410644
- Brooke, J. (1996). SUS: a "quick and dirty" usability scale. In: P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (Hrsg.): Usability Evaluation in Industry. London: Taylor and Francis.
- Laugwitz, B., Schrepp, M. & Held, T. (2008). Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (Ed.): USAB 2008, LNCS 5298, pp. 63-76.

To have and to hold during L2: Hands on wheel keeps drivers in lane while mind on road speeds response

Annika F.L. Larsson^{1*}, Ulrich Sander¹, Bobbie Seppelt²

¹ Veoneer Research, Wallentinsvägen 22, Vårgårda, Sweden

² MIT AgeLab, Cambridge MA, USA

*annika.fl.larsson@gmail.com

Abstract: This paper provides insights into response and recovery following silent system malfunctions when drivers either keep their hands on or off the steering wheel during automated lateral and longitudinal control (Level 2). A test track study was carried out with 37 participants, with a between-group design for hands on or off the steering wheel. In automated mode, every other lap was driven with/without cognitive load (implemented via a 1-back task). Participants experienced three unexpected steering events per lap requiring manual steering input. In the final lap, an unexpected lead vehicle event required drivers to brake or steer to avoid crashing. Maximum steering amplitude as well as the standard deviation of steering acceleration were larger for the hands-off group, without any effect of cognitive load. Hands-off participants also deviated further from lane center. For the unexpected lead vehicle event there was a trend that mind-on drivers, regardless of hands on/off, were quicker to brake than mind-off drivers.

1. Introduction

Recently, systems which explicitly allow the driver to remove their hands from the steering wheel while maintaining their eyes on the road have become commercially available. These hands-off (“L2”) assistance systems provide distance- and speed-keeping, and keep the vehicle within its lane. L2 systems still require drivers to look at the road and respond to any situation that could be risky or undesirable, making the physical driving task an intermittent activity with the potential of long periods of non-activity.

This paper addresses two main research questions related to driving with L2 automation while experiencing repeated silent steering failures: (i) How does hands on/off the steering wheel affect initial reaction and recovery characteristics in a repeated steering event and in a single critical lead vehicle event, and (ii) how are reaction and recovery characteristics to lateral and longitudinal events affected by working memory load.

1.1 Background

Garbacik et al., (2021) found that for unexpected steering events, only drivers who had at least one hand on the steering wheel managed to stay in lane. Without hands on the wheel, drivers lose the neuro/sensory connection to appropriately guide their response to unexpected events (Benderius & Markkula, 2014; Mole et al., 2019). Therefore, they are expected to exhibit miscalibrated steering responses, such as higher steering amplitude or steering acceleration (Navarro, François, & Mars, 2016). For steering quality, McDonald et al., (2019) emphasize the difference between the response and recovery phase, pointing out that different models are needed to account for action selection (response) and then post-takeover (recovery) performance.

Victor et al., (2018) conducted a test-track experiment on vehicle automation, where nearly 30% of hands-on eyes-on drivers still crashed into a stationary object. For such longitudinal events, research suggests that there is no difference in response for driving hands-on or hands-off

(Damböck, Weissgerber, Kienle, & Bengler, 2013). Victor et al. (ibid) suggest the reason for a lack of response was automation performing well for 30 minutes before the event. Drivers respond more quickly when they recognize the need for action (c.f. Larsson, Kircher, & Andersson Hultgren, 2014; Seppelt & Lee, 2019).

For the mental control loop, working memory load inhibits a driver's ability to assimilate cues and predict on-road events (e.g. Myers, Stokes, & Nobre, 2017). Predictive processing is essentially shut down (Engström, Markkula, Victor, & Merat, 2017), meaning that the top-down processes required to switch attention back to driving are unavailable. Drivers are thus only able to initiate automatized responses to events, such as responding to looming (Markkula, Engström, Lodin, Bårgman, & Victor, 2016).

Here too, this study seeks to understand the interacting effect of working memory load on automatized and controlled driver response behaviors.

2. Method

2.1 Participants

Participants were recruited via Facebook ads, and completed a screening questionnaire to balance the technology interest and age of the participants across groups. In total, 37 participants completed the drives with full vehicle data. The average and median age was 41 years (age range: 25-60). All participants drove a car on a daily basis.

2.2 Procedure

The study was conducted at the 5.7 km AstaZero rural road track in Sweden. Participants drove a Lincoln MKZ equipped with steering and speed keeping automation, based on GPS track following. The vehicle did not use any sensors for its automation, and longitudinal control was only paused (not deactivated) upon use of the brake pedal.

Table 1 Study design

| Hands On Wheel | | Hands Off Wheel | |
|----------------|---------------|-----------------|---------------|
| MindOn first | MindOff first | MindOn first | MindOff first |
| 9 | 8 | 10 | 10 |

37 participants in total

Driving 65 km/h, Participants completed an introductory lap of fully manual driving, followed by four in automated mode, with at least one car safely overtaking the test vehicle each lap. 17 participants drove all laps hands-on, and 20 participants drove all laps hands-off (see Table 1). In automated mode, every other lap was driven with/without cognitive load. Cognitive load was imposed by means of the 1-back task as described in Mehler, Reimer, & Dusek, (2011).

During each automated lap, participants experienced three different steering events that required manual steering input to avoid going off road, approximately once every two minutes. These events included both omission (not turning when it should have, three seconds) and commission (turning when it should not, 8 degrees, two seconds). Thereafter the vehicle reactivated the normal GPS path-following. The steering events were validated through a test drive with active steering system developers from Zenuity.

Steering events were carefully placed at safer sections of the test track, with the first (omission) event repeated every lap. A final event was designed to mimic that of Victor et al., (2018) where a lead vehicle did a cut-out in front of a stationary (soft) target vehicle. Automated steering was deactivated for safety reasons until the car had passed the target vehicle.

This paper reports only on the first (repeated) event of each lap and the final event with the soft vehicle target.

3. Results

3.1 Response and recovery metrics

For the first event, the maximum steering amplitude ($F(1,33) = 28.49, p < .001$) as well as the standard deviation of steering acceleration ($F(1,33) = 11.34, p < .01$) were larger for the hands-off group than the hands-on group, without any effect of mind on/off. For standard deviation of centerline offset, there was a clear difference of hands-off participants deviating further than hands-on ($F(1,33) = 57.50, p < .001$). There was also a trend ($F(1,33) = 3.02, p = .09$) that mind on/off influenced the standard deviation of centerline offset.

For the repeated events, there was a learning effect in maximum steering amplitude and in centerline offset from the first event for hands-off ($F(1,35) = 9.5, p < .001$; $F(1,35) = 7.80, p < .001$) but not for hands-on. In subsequent laps, there was no improvement in response quality but hands-off drivers

had higher maximum steering amplitude, $F(1,35) = 11.63, p < .01$) and higher centerline offset ($F(1,35) = 21.1, p < .001$), see Appendix A, Figure 1 and Figure 2.

For the soft vehicle target, the steering reaction time was no different between hands on/off. For brake response time, this event indicated a trend ($F(1,35) = 3.51, p = .07$) that mind-on drivers, regardless of hands on/off, were faster to initiate braking than mind-off drivers (Appendix A, Figure 3).

3.2 N-back performance

On average, the hands-off group outperformed the hands-on group at least for the first two laps. Performance improved after the first two laps for both groups. Two participants were excluded from analysis due to misunderstanding the n-back task. Due to the unbalanced nature of the study design, an ANOVA could not be performed.

4. Discussion

In repeated steering events, hands on or off the steering wheel affects response quality substantially for the first failure. The results for the hands-off drivers could be related to the expectation that the system was more competent, a misconception corrected by the first silent omission error. For subsequent failures, hands-off drivers continued to end up further outside their lane and with higher maximum steering amplitude, showing a miscalibration in their response process.

For action selection in the unexpected lead vehicle event, mind-off drivers tended to take longer, or they responded when looming had become evident (Markkula et al., 2016). Automation supervision, in essence, introduces cognitive control requirements back into automatized manual control behavior (Engström et al., 2017).

5. Conclusions

No participant in the present study went off the road or collided into an obstacle. This may be due to the early occurrence of a steering event, and that such events kept happening. However, our results indicate that imperfect automation systems may not be sufficient to address neither the neuro/sensory disconnect from having hands off the wheel (i.e., the disruption in automatized processing), nor the automation-induced need to consciously process whether events will require a deliberate (controlled) response. These results have implications for the design of L2 systems in how engaging drivers' hands and mind affect reactive and predictive driver behavior, respectively.

6. Acknowledgments

The authors would like to thank the team at Veoneer Research for their support and skill in performing the test track tests, and Dr Jonas Andersson at RISE for his contributions to the method development.

References

- Benderius, O., & Markkula, G. (2014). Evidence for a fundamental property of steering. *Proceedings of the Human Factors and Ergonomics Society, 2014-Janua*, 884–888. <https://doi.org/10.1177/1541931214581186>
- Damböck, D., Weissgerber, T., Kienle, M., & Bengler, K. (2013). Requirements for cooperative vehicle guidance.

IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, (Itsc), 1656–1661.
<https://doi.org/10.1109/ITSC.2013.6728467>

Engström, J., Markkula, G., Victor, T., & Merat, N. (2017). Effects of Cognitive Load on Driving Performance: The Cognitive Control Hypothesis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 59(5), 734–764. <https://doi.org/10.1177/0018720817690639>

Garbacik, N., Mastory, C., Nguyen, H., Yadav, S., Llaneras, R., & McCall, R. (2021). Lateral Controllability for Automated Driving (SAE Level 2 and Level 3 Automated Driving Systems). *SAE Technical Papers*, (2021), 1–10. <https://doi.org/10.4271/2021-01-0864>

Larsson, A. F. L., Kircher, K., & Andersson Hultgren, J. (2014). Learning from experience: Familiarity with ACC and responding to a cut-in situation in automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 46. <https://doi.org/10.1016/j.trf.2014.05.008>

Markkula, G., Engström, J., Lodin, J., Bärgrman, J., & Victor, T. (2016). *A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies. Accident Analysis and Prevention (Vol. 95).* <https://doi.org/10.1016/j.aap.2016.07.007>

McDonald, A. D., Alambeigi, H., Engström, J., Markkula, G., Vogelpohl, T., Dunne, J., & Yuma, N. (2019). Toward Computational Simulations of Behavior During Automated Driving Takeovers: A Review of the Empirical and Modeling Literatures. *Human Factors*, 61(4), 642–688. <https://doi.org/10.1177/0018720819829572>

Mehler, B., Reimer, B., & Dusek, J. A. (2011). *MIT AgeLab Delayed Digit Recall Task (n-back)*. MIT AgeLab White Paper Number 2011-3B. Massachusetts Institute of Technology, Cambridge, MA.

Mole, C. D., Lappi, O., Giles, O., Markkula, G., Mars, F., & Wilkie, R. M. (2019). Getting back into the loop : the perceptual-motor determinants of successful transitions out of automated driving. *Human Factors*, 61(7), 1037–1065.

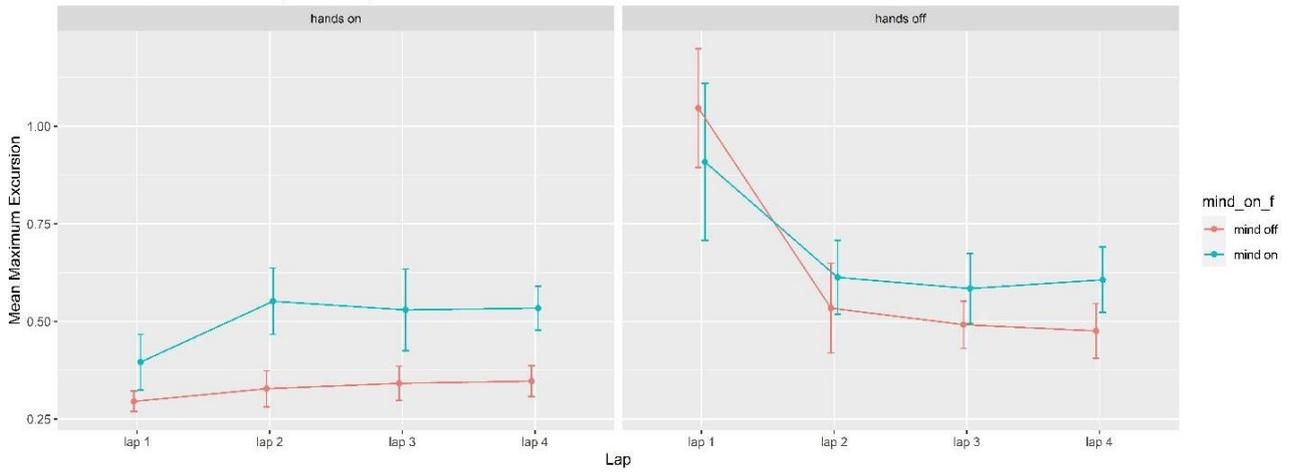
Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends in Cognitive Sciences*, 21(6), 449–461. <https://doi.org/10.1016/j.tics.2017.03.010>

Navarro, J., François, M., & Mars, F. (2016). Obstacle avoidance under automated steering: impact on driving and gaze behaviours. *Transportation Research Part F: Traffic Psychology and Behaviour*. <https://doi.org/10.1016/j.trf.2016.09.007>

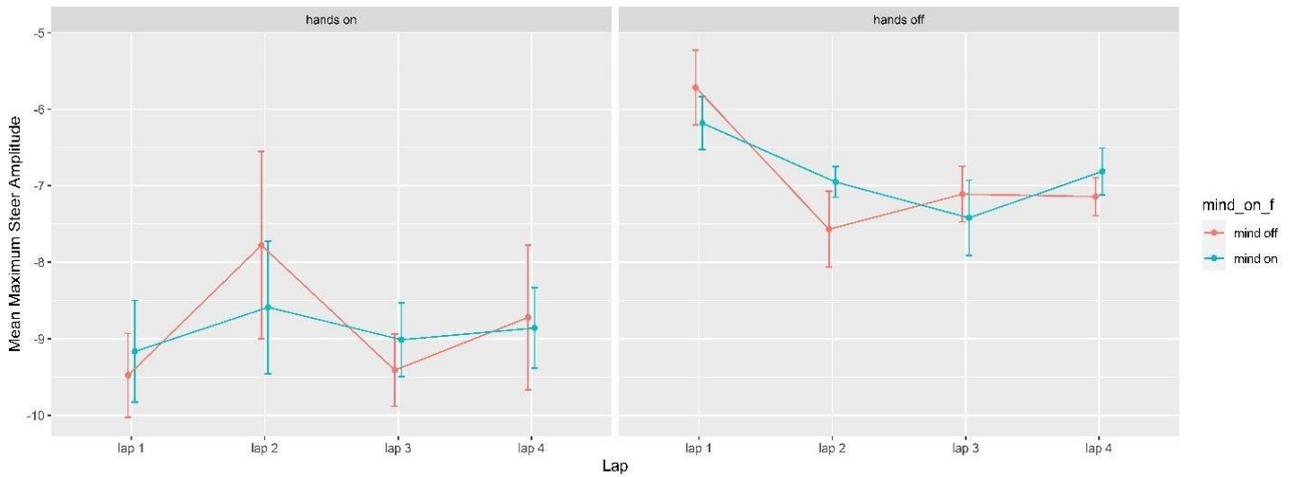
Seppelt, B. D., & Lee, J. D. (2019). Keeping the driver in the loop: Dynamic feedback to support appropriate use of imperfect vehicle control automation. *International Journal of Human Computer Studies*, 125(608), 66–80. <https://doi.org/10.1016/j.ijhcs.2018.12.009>

Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human Factors*, 60(8), 1095–1116. <https://doi.org/10.1177/0018720818788164>

Appendix A

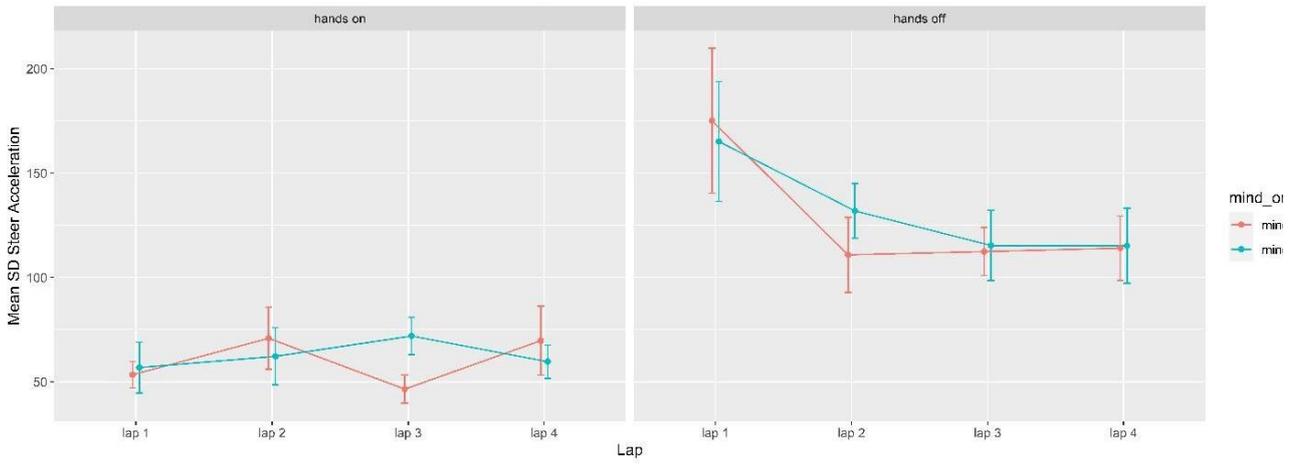


(a)

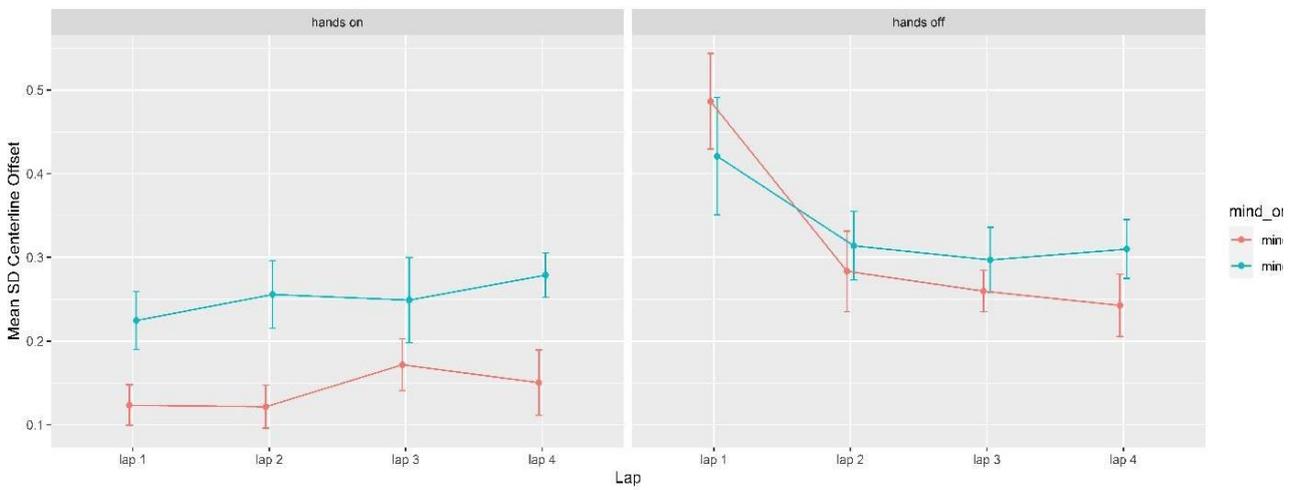


(b)

Fig. 1. Driver first response for the repeated steering event, **(a)** Mean maximum excursion, **(b)** Mean maximum steering amplitude



(a)



(b)

Fig. 2. Driver recovery for the repeated steering event, (a) Mean standard deviation of steering acceleration, (b) Mean standard deviation of centerline offset

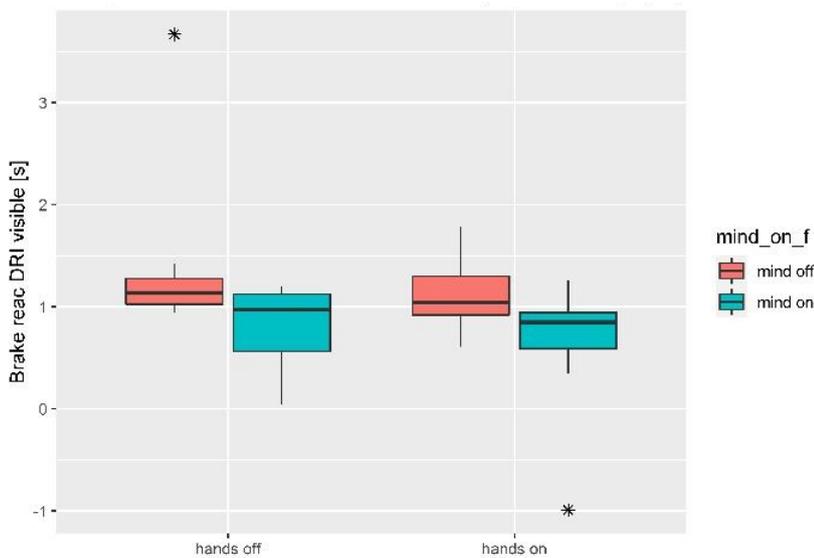


Fig. 3. Brake reaction time to the soft target vehicle

Attentional demands of using an application for real-time traffic information feedback in road operators' vehicles

Lara Désiré^{1*}

¹ Cerema, Research team PsyCAP, 5 rue Jules Vallès, F-22015, Saint-Brieuc, France

*lara.desire@cerema.fr

Abstract: Using cooperative ITS services such as a traffic information feedback application could help road operators to inform road users in real time about traffic conditions. However, the interaction tasks involved should not distract road operators. We studied the effects of three levels of interaction task complexity using an application currently being deployed. 18 participants completed different tasks while driving an instrumented vehicle. Depending on the availability of the event to be reported on the first or second screen of the application, the interaction tasks required 1 press (first screen) vs. 2 presses (top of second screen) vs. 1 press, drag and drop and 1 press (bottom of second screen). The subjective, temporal, visual and cognitive demands of the three application interaction tasks were compared with each other and with a radio task deemed acceptable, and with two “artificial” tasks recognised for their high-level visual or cognitive demands. Subjective, temporal, visual and cognitive demands were higher for events reported from the second screen vs. the first screen. Compared to the high-demand tasks and the radio task, the interaction tasks with the first screen required lower subjective, temporal and cognitive demands and a similar visual demand. These findings suggest that the use of a traffic reporting application by road operators requiring no more than one press would be feasible without undue distraction, but that more complex interactions could pose a risk in the absence of training to lower the attentional demand.

1. Introduction

Thanks to the deployment of cooperative ITS (C-ITS), road operators can alert road users in real time of their interventions on the road network and the events they observe. This could help to improve comfort and road safety for both road operators and road users (AIPCR, 2019). As part of the European SCOOP project for the pre-deployment of C-ITS, the study of the *a priori* acceptability of the deployment of a traffic information reporting application (SCOOP application) among French road operator agencies revealed that one of the obstacles concerned the interference with the driving activity of road operators (Chahir et al., 2019).

In the scope of the ensuing C-Roads European project, this potential obstacle prompted the need to consider the distraction induced by the traffic information reporting application among road operators during driving. In order to investigate the distraction issue, a study was conducted to evaluate the intensity of the attentional demand related to different interaction tasks with the SCOOP application in order to help decision-making on the work processes of road operators with this application.

The methodology used was based on those developed to assess the impact of different types of interactions with in-vehicle technologies (Mehler et al., 2016; Reimer et al., 2013; Strayer et al., 2019; Zhang et al., 2015). The objective was to compare the demand required by three different levels of interaction task complexity with the application in terms of subjective, temporal, visual or cognitive demands (Strayer et al., 2019).

2. Method

2.1 Participants

18 participants (all males) were recruited by one of the road operator agencies in which the SCOOP application was under deployment. Considering that it was conducted in only one of its departments, we tried to balance the study's participants among 3 different professional groups with potentially different knowledge of the application.

2.2 Equipment, driving route and procedure

Participants drove an instrumented vehicle equipped with a digital tablet and, in particular, 4 cameras, an eye-tracker and a Detection Response Task (DRT) kit configured with the vibro-tactile stimulus (ISO, 2016).

The experimental route was a two-lane dual carriageway with a speed limit of 110 km/h on the major portion. The procedure consisted of one familiarisation and 5 experimental sessions. Each experimental session began with a session to train the participants on the different tasks, and ended with their completion of the perceived mental workload scale (Reimer et al., 2013).

2.3 Tasks

The application interaction tasks consisted in reporting an event supposedly encountered on the road while driving. The required interactions involved three different levels of interaction complexity: 1) an event available on the first screen, requiring 1 press; 2) an event available on the top of the second screen, requiring 2 presses, and 3) an event available on the bottom of the second screen, requiring 1 press, a drag and drop then 1 press.

These tasks were compared with a given acceptable radio task as specified by the automotive industry (Alliance of Automobile Manufacturers & Driver Focus-Telematics Working Group, 2006). Each participant also performed two secondary "artificial" reference tasks recognised for their high visual (Surrogate Reference Task) and cognitive (2-back task) demands (Mehler et al., 2011; Reimer et al., 2013; Zhang et al., 2015). These tasks were used to calculate standardised scores.

2.4 Data and statistical analysis

For each trial involving the different tasks, the following items of data were recorded: 1) perceived mental workload, 2) time to complete the task, 3) percentage of time spent looking at the road while performing the task, and 4) reaction time to the DRT.

These data items were respectively used to calculate four standardised scores (Strayer et al., 2019): 1) subjective demand, 2) temporal demand, 3) visual demand and 4) subjective demand.

For each score, 3 different mixed models (Task type; Group; Task type x Group) were compared using the Bayesian Information Criteria. In all cases, the most likely model was the model considering only the Task type.

3. Results

The Task type had a significant effect on the four standardised scores (Table 1).

The subjective demand and task duration time were higher for interactions with the second screen than the first screen. Furthermore, they were higher for interactions with the bottom than the top of the second screen. Also, the mean subjective demand and task duration were below the high references for any of the interactions with the application (Fig. 1 a & b).

Visual and cognitive demands were higher for interactions with the second screen than the first screen, but similar for interactions with the top and the bottom of the second screen. Furthermore, the mean visual and cognitive demands were above the high references for any of the interactions with second screen of the application (Fig. 1 c & d).

Compared to the radio task, the interaction task with the first screen had lower subjective, temporal and cognitive demands but a similar visual demand. The interaction tasks

with either the top or bottom of the second screen had a lower temporal demand, a similar cognitive demand and a higher visual demand. The subjective demand was similar for the top, but tended to be higher for the bottom of the second screen.

4. Discussion

The four indices of attentional demand show that the demand was higher when reporting an event from the second screen than from the first screen of the traffic information feedback application.

Concerning the first screen, the subjective, temporal and cognitive demands were lower than those observed for the radio and the high-demand tasks, and the visual demand was similar to that observed for the radio task and high-demand tasks. The application's use thus seems possible for events accessible from this first screen.

Concerning the second screen, the temporal demand was lower than the high reference threshold set and that observed for the radio task. The subjective demand was also lower than the high reference threshold, and similar to that observed for the radio task. Nevertheless, the visual demand was higher than that observed for the radio and the high visual demand tasks. For the cognitive demand, it was similar to that observed for the radio and the high cognitive demand tasks. Consequently, with the current state of the interface, the application's use to report events from the second screen does not seem possible while driving.

5. Conclusions

The use of a traffic information feedback application by road operators seems possible without compromising their safety under certain conditions. An application used to report an event from a first screen presenting 5 items requires a low to moderate attentional demand, compatible with the driving activity. Nevertheless, the need to switch to a second screen to choose among different items and to possibly scroll through them requires a high attentional demand likely to cause distraction. This result calls for particular vigilance when training road operators in the application's use as a means to improve their knowledge of the location of events in order to reduce the attentional demand required for reporting events that are not directly accessible.

Table 1 Results of the statistical analysis for each score with a mixed model considering the Task type.

| Demand standardised score | SS | MS | NumDF | DenDF | F | p |
|---------------------------|--------|---------|-------|--------|--------|---------|
| Subjective | 5.1179 | 1.706 | 3 | 51 | 17.301 | < 0.001 |
| Temporal | 26.821 | 8.9405 | 3 | 232.51 | 73.069 | < 0.001 |
| Visual | 1.6766 | 0.55887 | 3 | 218.25 | 8.3614 | < 0.001 |
| Cognitive | 419.54 | 139.85 | 3 | 306.26 | 6.7692 | < 0.001 |

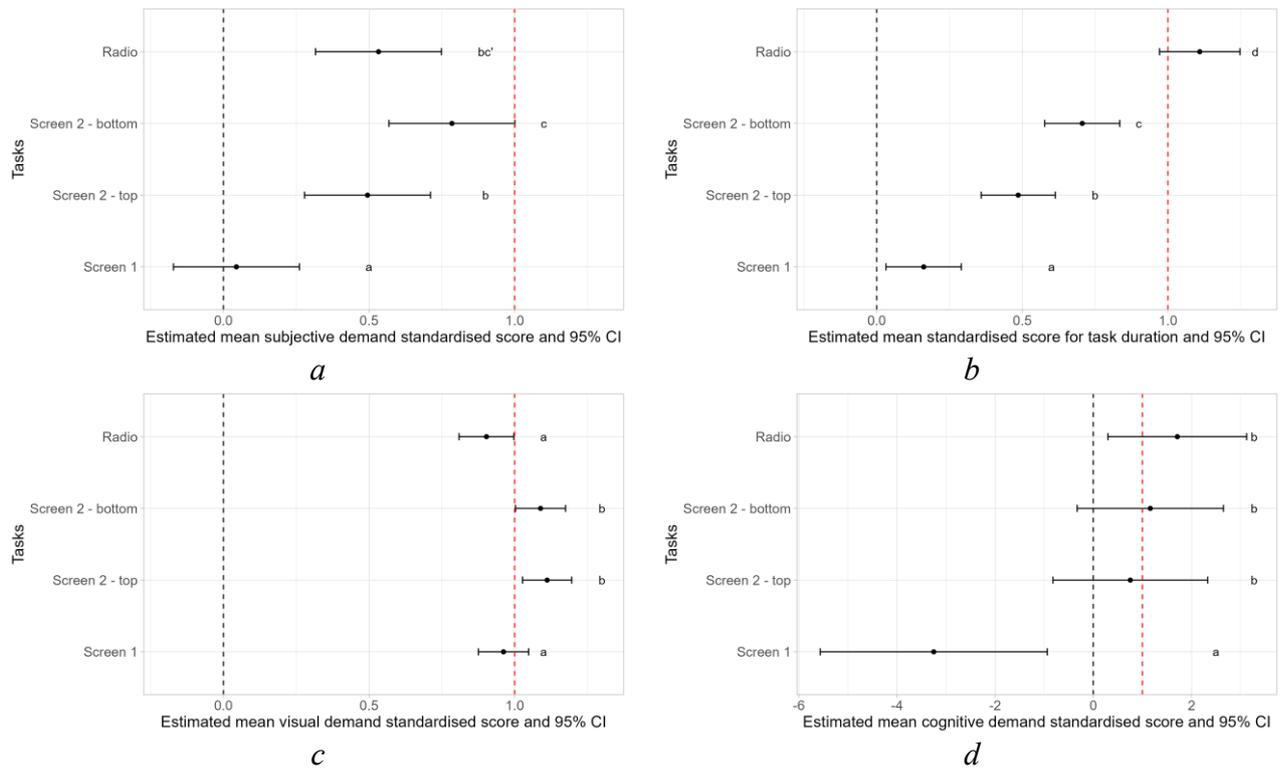


Fig. 1. Estimated mean standardised score of the attentional demand and 95 % confidence interval as a function of task type. The dashed vertical black lines represent the score for the single task performance. The dashed vertical red lines represent the score for the high demand reference tasks or for a duration of 24 seconds. Scores with different letters are significantly different ($p < 0.05$). Scores with similar letters and a quotation mark are not significantly different but approach significance ($p < 0.1$).
(a) Subjective, **(b)** temporal, **(c)** visual, and **(d)** cognitive demands.

6. Acknowledgments

Funding for this study was provided in part through the European-funded C-Roads France project. The views and conclusions expressed are those of the author and may not necessarily reflect those of the consortium members.

Many thanks to R. Gritti for his assistance with the experiment and C. Bourgeois for her participation in the pre-tests. I thank N. Chebassier, P. Ecobichon, N. Le Goff, L. Lilas, G. Morvan, F. Talbourdet (DIR Ouest) and Y. Le Chanu (Viveris) for their assistance in recruiting participants or preparing the app. I thank X. Bertrand and H. Poirier (Cerema, D2PN) for the implementation of the Surrogate Reference Task application. Thank you to G. Bongers (Noldus IT), U. Löfberg (SmartEye), J. Cooper (Red Scientific), N. du Lac (Intempora) and their colleagues for their support and responsiveness, and for making the best of their respective tools. Special thanks to A. Bousquié for creating the icons. And I would like to thank BIRDWELL Translations for their linguistic support.

References

AIPCR. (2019). *Connected vehicles—Challenges and opportunities for road operators* (N° 2019R11EN; p. 84). AIPCR. [https://www.piarc.org/fr/fiche-](https://www.piarc.org/fr/fiche-publication/30786-fr-Les_v%C3%A9hicules_connect%C3%A9s_D%C3%A9fis_et_opportunit%C3%A9s_pour_les_op%C3%A9rateurs_routiers)

publication/30786-fr-Les véhicules connectés - Défis et opportunités pour les opérateurs routiers Alliance of Automobile Manufacturers, & Driver Focus-Telematics Working Group. (2006). *Statement of principles, criteria and verification procedures on driver interactions with advanced in-vehicle information and communication systems including*. <https://autoalliance.org/wp-content/uploads/2018/08/Alliance-DF-T-Guidelines-Inc-2006-Updates.pdf>

Chahir, M., Bordel, S., & Somat, A. (2019). *Étude d'impact organisationnel chez les gestionnaires routiers—Rapport final de l'action de recherche—Accompagnement des évolutions organisationnelles et humaines introduites par la conception et le déploiement d'une innovation technologique (C-ITS) par l'analyse du système humain-technologie-organisation* (Livrable du projet SCOOP 2.3.5.3; p. 94).

ISO. (2016). *ISO 17488:2016—Road vehicles—Transport information and control systems—Detection-response task (DRT) for assessing attentional effects of cognitive load in driving* (ISO 17488:2016). AFNOR.

Mehler, B., Kidd, D., Reimer, B., Reagan, I., Dobres, J., & McCart, A. (2016). Multi-modal assessment of on-

- road demand of voice and manual phone calling and voice navigation entry across two embedded vehicle systems. *Ergonomics*, 59(3), 344-367. <https://doi.org/10/gf7972>
- Mehler, B., Reimer, B., & Dusek, J. A. (2011). *MIT AgeLab delayed digit recall task (n-back)* (MIT AgeLab White Paper N° 2011-3B; p. 22). Massachusetts Institute of Technology. http://agelab.mit.edu/system/files/2018-12/Mehler_et_al_n-back-white-paper_2011_B.pdf
- Reimer, B., Mehler, B., Dobres, J., & Coughlin, J. F. (2013). *The effects of a production level « voice-command » interface on driver behavior : Reported workload, physiology, visual attention, and driving performance* (MIT AgeLab Technical Report N° 2013-17A). Massachusetts Institute of Technology. https://www.researchgate.net/publication/258757705_The_Effects_of_a_Production_Level_Voice-Command_Interface_on_Driver_Behavior_Reported_Workload_Physiology_Visual_Attention_and_Driving_Performance
- Strayer, D. L., Cooper, J. M., Goethe, R. M., McCarty, M. M., Getty, D. J., & Biondi, F. (2019). Assessing the visual and cognitive demands of in-vehicle information systems. *Cognitive Research: Principles and Implications*, 4(1), 18. <https://doi.org/10/gf7kgd>
- Zhang, Y., Angell, L., Pala, S., & Shimonomoto, I. (2015). Bench-marking drivers' visual and cognitive demands : A feasibility study. *SAE International Journal of Passenger Cars - Mechanical Systems*, 8(2), 584-593. <https://doi.org/10/gf85rh>

Human-machine interface designs assisting drivers of automated vehicles during transitions: evaluation from an end-user perspective

Arun Muthumani^{1*}, Da Wang¹

¹ Autoliv Research, Wallentinsvägen 22, Vårgårda, Sweden

* arun.muthumani@autoliv.com

Abstract: Vehicles operating with multiple levels of automation include transitions: procedures changing vehicle automation state. An efficient human-machine interface (HMI) is needed to deliver information to drivers during state changes. In this study, three different HMI designs (Baseline, HMI-1, and HMI-2) to communicate upcoming transitions (to and from the vehicle) were developed based on the HMI framework method. All three HMI designs provide multimodal cues: a combination of visual, auditory, and haptic cues. HMI-1 includes an additional visual cue positioned on the circumference of steering wheel while HMI-2 includes a visual cue on the windshield. The HMI designs were evaluated in a driving simulator experiment with 24 participants. Both critical (unplanned) and non-critical (planned) transitions were presented to distracted (by a game on a touch display) participants at different time intervals. System usability (SUS) scores, a subjective measure, were higher for HMI-1 (mean score 81) than for two designs (Baseline:79, HMI-2: 79), but differences were not significant. HMI-1 had most desirable aspects. In a preference ranking, 67% of participants choose HMI-1 first, indicating higher desirability than both other designs which were each ranked first by only 17%. User comments received in post-experiment interviews indicated the desire for personalized HMIs to increase acceptance. From this study we conclude a preference for the HMI-1 design and recommend personalized HMIs with visual cues on the circumference of the steering wheel to be further developed and implemented in vehicles with multiple levels of automation.

1. Introduction

Vehicles equipped with multiple automation levels need to effectively, communicate their intentions and limitations to drivers, creating a need for a novel Human-Machine interface (HMI). Previous studies have experimented using multimodal HMIs (visual, auditory, haptic etc.) to communicate automation related information as multimodality improves recognition, understanding and promotes faster (van Erp et al., 2015; Petermeijer et al., 2016) and intuitive interactions with users (Manawadu et al., 2017). Studies on transitions used light displays such as LED strips mounted on the windshield (Yang et al., 2018) steering wheel (Muthumani et al., 2020) door panel (Wilbrink et al., 2020) or even entire vehicle interiors (Dziennus et al., 2016) for communication. However, only few studies investigated which position is best suited. One such recent study (Feierle et al., 2020) compared two positions (steering wheel vs windshield) for the additional light display with an auditory cue and found no significant difference for objective takeover time and subjective ratings. However, the study mainly considered two modalities for their investigation. Additionally, the study did not report on end-users' preference towards HMI solutions.

1.1 Objectives

This study aims to compare the additional visual displays mounted at two positions (steering wheel vs windshield) assisting drivers when shifting from one automation level (AD-Assisted mode) to another (PD- Piloted mode). For this purpose, three multimodal HMIs with (HMI-1 and HMI-2) and without (baseline) additional light display informing drivers about upcoming transition were developed. The

automation levels and the use cases used for the evaluation were based on the concepts in the Mediator project (Christoph et al., 2019). The objective of this study is to evaluate all three HMI designs in terms of perceived usability, desirability, and user preferences.

2. Method

2.1 Participants

Twenty-four volunteers (2 female and 22 male) who were employees and consultants working at different departments at Autoliv, participated in the study with mean age of 44 years (SD=11.5), holding a valid driving license with driving experience ranging from 1-47 years.

2.2 Human-machine interface designs

In this study, multimodal HMI feedback systems were developed. The interaction between the various HMI components and driver were designed using the HMI framework method (Diederichs et al., 2020). Figure.1 represents an example of HMI framework developed for the "HMI-1" design. The baseline design uses, instrument cluster (for visual icons) in-vehicle speakers (for sounds and voice messages) and seat belt (for belt pull). The other two designs provide an additional visual cue generated by 33 LEDs positioned of the circumference of steering wheel (for HMI-1) and an array of 20 LEDs mounted on the windshield (for HMI-2). Additionally, the touchpad interface positioned on steering wheel yoke (both left and right) are used for activation of automation modes. In the baseline design, AD and PD mode availability is conveyed by a notification sound, voice message along with an animated

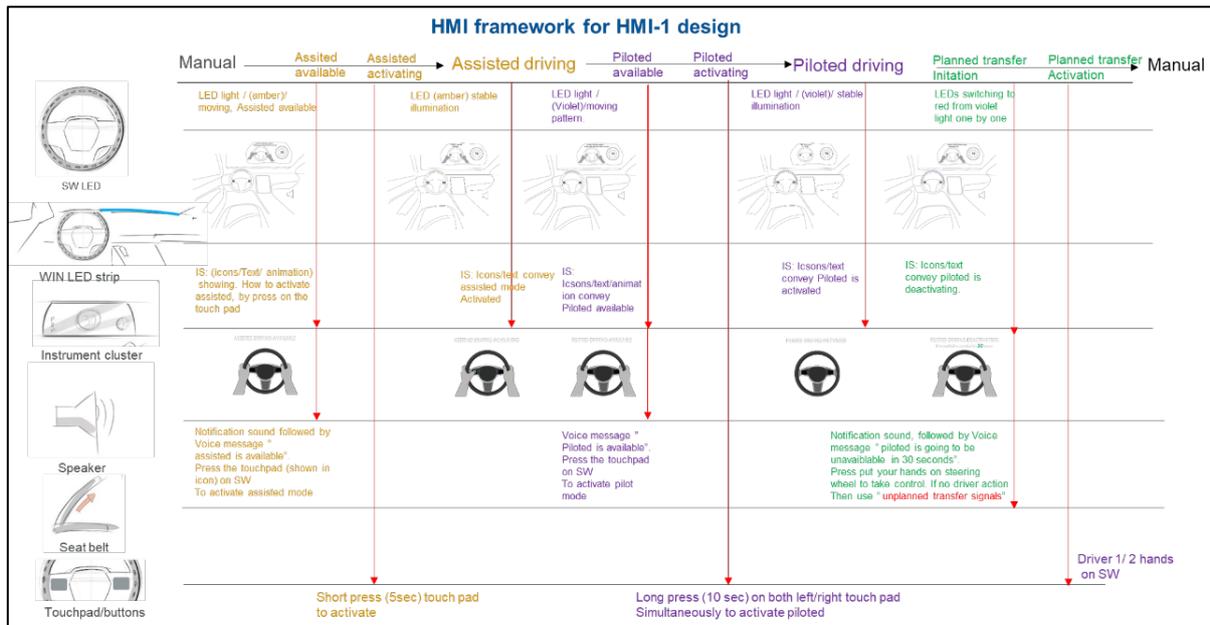


Fig. 1. HMI framework method for HMI-1 design

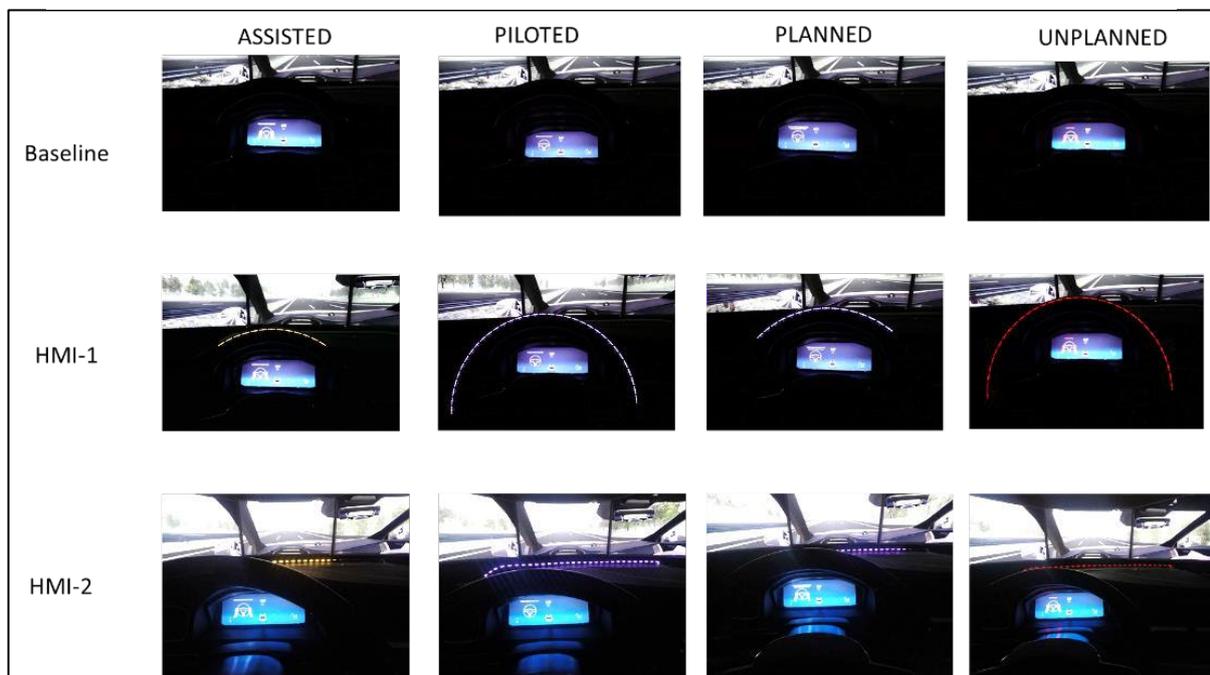


Fig. 2. All three HMI designs, automation levels and transitions

icon. The successful activation is conveyed with similar set of cues. The planned transition (from PD to manual) is communicated by a voice message followed by a “clock ticking sound” along with pulsating “hand-on wheel” icon including text display “Piloted driving disengaging, please take control” and a countdown timer. During unplanned transition (from PD to manual), the vehicle delivers a recurring voice message “Please take control”, continuous beeps” and a pulsating “hands-on wheel” icon with a display text (see Fig.2). Both HMI-1 and HMI-2 uses all set of cues from baseline design. However, the light pattern of the additional display is differentiated between the two designs (see Fig.2). In planned transitions LEDs switch off successively in a synchronised manner with “clock ticking

Table 1 Experiment order

| Groups | Lap 1 | Lap 2 | Lap3 |
|--------|------------------------|-----------------------|------------------------|
| G1 | Bas-(P)→(U) (P)→(U) | H1-(U)→(P) (P)→(U) | H2-(P)→(U) (P)→(U) |
| G2 | H2-(P)→(U) (P)→(U) | H1-(U)→(P) (P)→(U) | Bas-(P)→(U) (P)→(U) |
| G3 | Bas-(P)→(U) (P)→(U) | H2-(U)→(P) (P)→(U) | H1-(P)→(U) (P)→(U) |
| G4 | H1-(P)→(U) (P)→(U) | H2-(U)→(P) (P)→(U) | Bas-(P)→(U) (P)→(U) |

Bas-Baseline design; H1: HMI-1 design; H2: HMI-2 design; P: Planned transition; U: Unplanned transition

sound. During unplanned transitions LEDs start to pulsate in red.

2.3 Experimental design and procedure

A fixed base driving simulator was used for the study. The simulated driving scenarios consists of a three-lane motorway with no traffic. In AD mode, participants were asked to release their feet from vehicle controls (gas and brake pedal), but to keep the hands on the steering wheel without any steering. In PD mode, participants were asked to take their hands off the steering wheel and play a digital game (DOTS) on the centre stack display until receiving a transition request from the vehicle. Participants were divided into four groups (see Table. 1) and drove three laps (16 mins per lap) with one HMI design. At the end of each lap, participants were asked to fill in both a system usability score (SUS) (Jordan et al., 1996) and HMI desirable aspects (Richardson et al., 2018) questionnaire. After completing all three laps, participants were asked to fill in the HMI design ranking questionnaire (custom-made for this study) followed by an interview.

3. Results

3.1 Usability: SUS score

HMI-1 received a higher mean SUS rating (81) than HMI-2 (79) and baseline (79), (see Fig.3). However, no statistical significance is observed in repeated-measure ANOVA between the HMI designs ($F(2,69)=0.19, p = 0.834$).

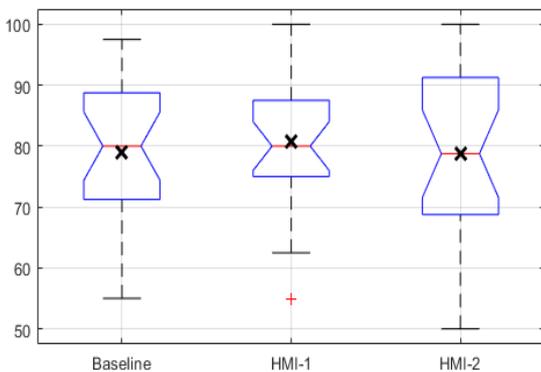


Fig. 3. SUS scores for all three HMI designs, x show the mean value, the red line is the median

3.2 HMI desirable aspects

Desirability of HMI designs is shown in six categories (see Fig.5). HMI-1 received the highest scores. For all items participants chose “strongly agree” more often for HMI-1 than for the other designs. For example, nearly 25% of participants strongly agreed to “provides appropriate trust” for HMI-1 design while only 8.3% strongly agreed to this with baseline and 16.7% with HMI-2.

3.3 HMI design ranking

Results from the HMI design ranking (see Fig.4), showed that 67% of participants have chosen HMI-1 as Rank 1, compared 17% for Baseline and 17% for HMI-2 design.

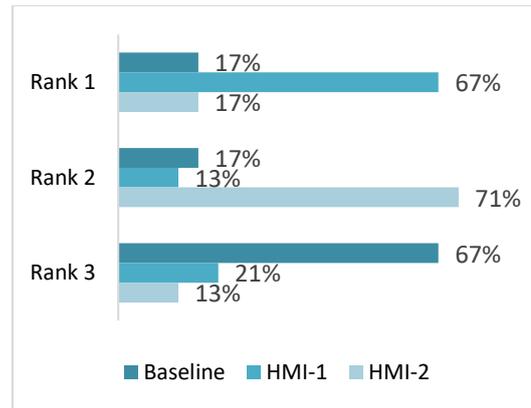


Fig. 4. Ranking for all three HMI designs

3.4 Interview

Unstructured participant feedback collected at the end of the experiment highlighted the importance of personalization HMI designs with user comments like: “I do like the LEDs showing the activation I like the way the LEDs progress towards taking control. Not sure if I really like the LEDs indicating the two modes are ready for activation. I think I would rather leave that function behind” Additionally, comments were also received on each HMI interfaces.

4. Discussion

Three HMI designs assisting drivers of automated vehicles during transitions were evaluated. HMI-1 included visual cues on the steering wheel and obtained the highest SUS score (81): this can be interpreted as “excellent” design according to the adjective rating scale (Bangor et al., 2009). However, no statistical difference is observed between the HMI designs which is in line with the study comparing additional displays (Feierle et al., 2020).

Additionally, user comments suggest refining the design according to individual preferences, i.e., personalization. Still, a proper balance between safety and user needs must be maintained.

There are limitations. Transition scenarios were investigated without presenting a hazard or threat situation in the simulated environment, which could influence the results. Additionally, LED illumination patterns were evaluated only in a simulator environment which could provide different user response when tested in a real vehicle. Another limitation is the sample size used for study. Future research investigation should focus on understanding the HMI design implications in a real vehicle environment and conducting more user tests to understand the personalization aspects in detail. As this could help HMI designer to understand balance between safety and user preferences.

The findings from this study contribute to designing the final HMI for the Mediator project (<https://mediatorproject.eu/>). The developed final HMI will further be evaluated with end-users to measure the effectiveness and user acceptance.

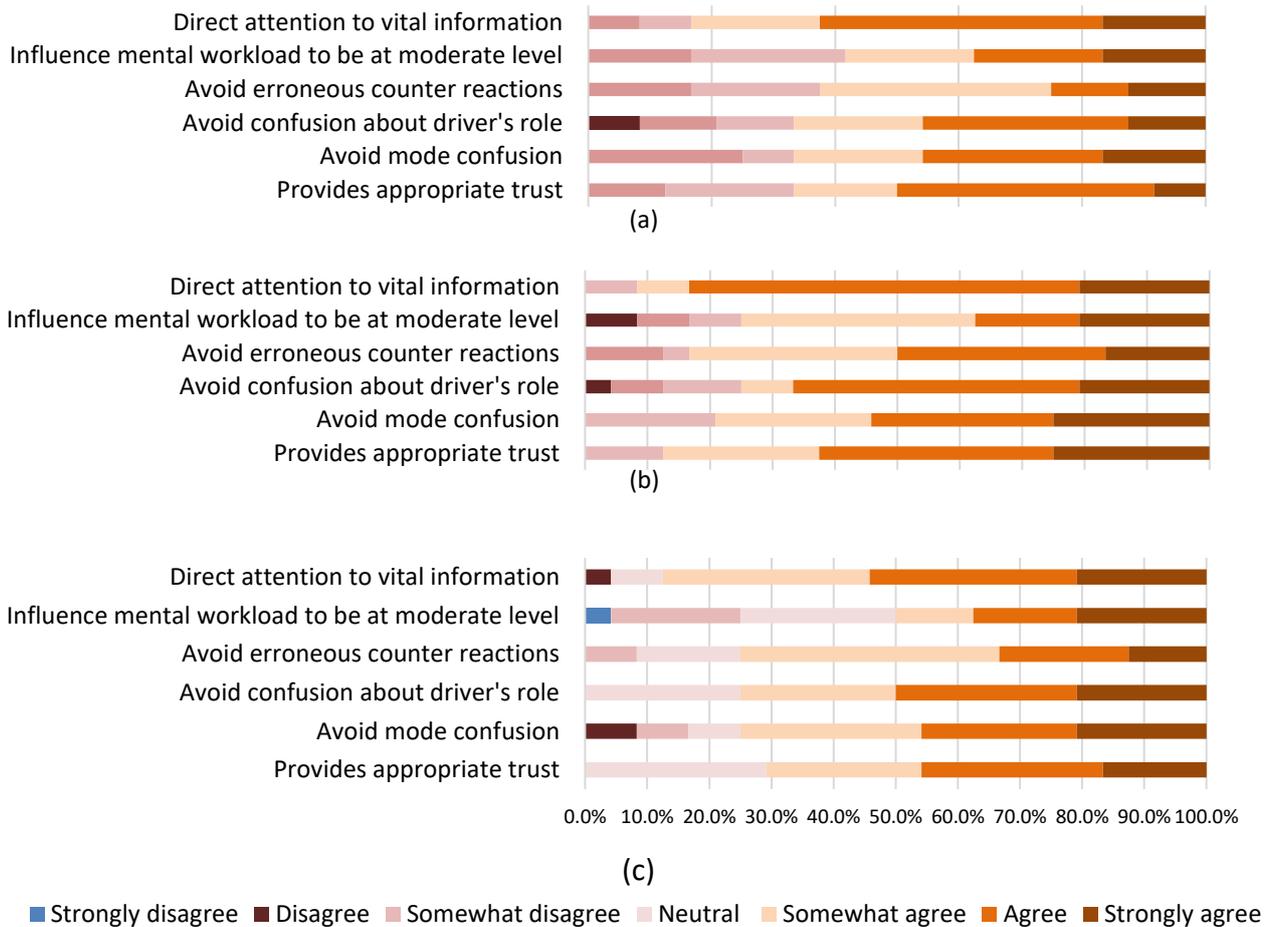


Fig. 5. HMI desirable aspects (a) Baseline, (b) HMI-1 and (c) HMI-2

5. Conclusions

HMI designs for mode transitions with visual cues on steering wheel improve perceived usability, desirability, and user preference. Enabling personalization of HMI elements can further improve the attractiveness of HMIs. We recommend personalized HMIs with visual cues on circumference of the steering wheel to be further developed and implemented in vehicles with multiple levels of automation.

6. Acknowledgments

This research was part of the Mediator project funded from the European Union's Horizon 2020 research programme under the grant agreement no 814735.

References

- Bangor, A., Kortum, P., and Miller, J. (2009). Determining what individual SUS Scores Mean: Adding Adjective rating scale for highly automated truck driving. *2009 Journal of usability studies* (Vol. 4, no.3, pp. 114–123).
- Christoph, M., Cleij, D., Ahlström, C., Bakker, B., Beggiato, M., Borowsky, A., ... Van Nes, C.N (2019). Mediating between human driver and automation: state-of-the art and knowledge gaps. D1.1 of the H2020 project MEDIATOR.
- Diederichs, F., Knauss, A., Wilbrink, M., Lilis, Y., Chrysochoou, E., Anund, A., Bekiaris, E., Nikolaou, S., Finér, S., Zanovello, L., Maroudis, P., Krupenia, S., Absér, A., Dimokas, N., Apoy, C., Karlsson, J., Larsson, A., Zidianakis, E., Efa, A., ... Bischoff, S. (2020). Adaptive transitions for automation in cars, trucks, buses and motorcycles. *IET Intelligent Transport Systems*, 14(8), 889–899. <https://doi.org/10.1049/iet-its.2018.5342>
- Feierle, A., Holderied, M., & Bengler, K. (2020). Evaluation of Ambient Light Displays for Requests to Intervene and Minimal Risk Maneuvers in Highly Automated Urban

Driving. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 1–8.
<https://doi.org/10.1109/ITSC45102.2020.9294645>

Jordan, P. W., Thomas, B., McClelland, I. L., & Weerdmeester, B. (Eds.). (1996). *Usability Evaluation in Industry* (0 ed.). CRC Press.
<https://doi.org/10.1201/9781498710411>

Manawadu, U. E., Kamezaki, M., Ishikawa, M., Kawano, T., & Sugano, S. (2017). A multimodal human-machine interface enabling situation-adaptive control inputs for highly automated vehicles. *2017 IEEE Intelligent Vehicles Symposium (IV)*, 1195–1200.
<https://doi.org/10.1109/IVS.2017.7995875>

Muthumani, A., Diederichs, F., Galle, M., Schmid-Lorch, S., Forsberg, C., Widroither, H., Feierle, A., & Bengler, K. (2020). How Visual Cues on Steering Wheel Improve Users' Trust, Experience, and Acceptance in Automated Vehicles. In N. Stanton (Ed.), *Advances in Human Aspects of Transportation* (Vol. 1212, pp. 186–192). Springer International Publishing. https://doi.org/10.1007/978-3-030-50943-9_24

Petermeijer, S. M., de Winter, J. C. F., & Bengler, K. J. (2016). Vibrotactile Displays: A Survey with a View on Highly Automated Driving. *IEEE Transactions on Intelligent Transportation Systems*, *17*(4), 897–907.
<https://doi.org/10.1109/TITS.2015.2494873>

Richardson, N. T., Lehmer, C., Lienkamp, M., & Michel, B. (2018). Conceptual design and evaluation of a human machine interface for highly automated truck driving. *2018 IEEE Intelligent Vehicles Symposium (IV)*, 2072–2077.
<https://doi.org/10.1109/IVS.2018.8500520>

van Erp, J. B. F., Toet, A., & Janssen, J. B. (2015). Uni-, bi- and tri-modal warning signals: Effects of temporal parameters and sensory modality on perceived urgency. *Safety Science*, *72*, 1–8.
<https://doi.org/10.1016/j.ssci.2014.07.022>

Psychosocial Needs and Factors Contribute to Problematic Phone Use while Driving

Yi-Ching Lee^{1*}, Chaur-Shyan Lee²

¹ Department of Psychology, George Mason University, Fairfax, VA, USA

² Department of Applied Economics, National Chung Hsing University, Taichung, Taiwan

* YLEE65@gmu.edu

Abstract: Cell phones (or smartphones) have become an essential part of daily functions. However, using the phone while operating machinery, such as driving, leads to problematic consequences and compromises safety. There have been legislative, technological, and educational strategies and countermeasures but their effectiveness on reducing distracted driving behaviours is inconclusive, suggesting the need to investigate other contributing factors and mitigation efforts. This article aims to propose a new framework that includes social needs, habitual use of smartphones, and self-control as core contributing factors to problematic phone use while driving. Preliminary data was collected and structural equation modelling is being used to test the relationships of the latent constructs within the framework. Results will be included soon. This framework bridges the distracted driving literature with psychological needs and technological adoption literature and provides testable hypotheses for future research. In addition, it offers suggestions for the framing of messaging for awareness champions and pledges as well as comprehensive research efforts to better understand the underlying contributing factors of problematic phone use while driving and distracted driving.

1. Introduction

Mobile cellular phones, especially smartphones (henceforth, cell phones), have become an integral part of our daily functions. The cell phone subscriptions have been growing steadily in the past two decades: in 2019 there were 109.46 subscriptions per 100 people worldwide and 134.46 subscriptions per 100 people in the U.S. (World Bank, 2019). Recent studies indicate that an average individual spends over four hours a day on their phone (Curtin, 2018) and 79% of 18 to 44 year-olds have their phone on them for 22 hours a day (Stadd, 2013).

What is alarming is when using cell phones while operating machinery, such as driving a car and being distracted by interactions with phones (Strayer & Fisher, 2016; World Health Organization & National Highway Traffic Safety Administration, 2011). Prior studies have documented the impact of visual, cognitive, and manual distraction and researchers have attempted to detect and mitigate distraction by technological means (Caird et al., 2014; Kashevnik et al., 2021; Leipnitz et al., 2022). For example, cell phone blocking technologies can be activated when built-in phone sensors detect movement such as in a moving vehicle and subsequently silence the phone, redirect incoming calls to voicemail, or automatically reply to incoming text messages with a pre-programmed message. Although a recent study indicated feasibility to use smartphone-based blocking technology to collect phone use data (McDonald et al., 2019), subjective data indicated that many drivers are motivated to find ways to cheat the technology or use a passenger's phone when theirs are locked while driving (Creaser et al., 2015).

In addition, legislative efforts have been introduced to regulate the use of cell phones and other electronics by drivers. Currently, 23 states, the District of Columbia (D.C.), Puerto Rico, Guam, and the U.S. Virgin Islands ban all drivers from using hand-held cell phones while driving. While some

studies document a limited or short-lived positive impact of the laws on cell phone use rate and traffic fatalities (Lim & Chi, 2013; Rocco & Sampaio, 2016), others find no evidence of change after legislative effects are in place (Highway Loss Data Institute, 2009; McCart & Geary, 2004)

Numerous local, state, and federal agencies are dedicated to end distracted driving by promoting campaigns (e.g., NHTSA's "U Drive. U Text. U Pay.") and increase public's awareness by issuing public service announcements. Additionally, cell phone service providers, car manufacturers, and insurance companies (e.g., AT&T's "It Can Wait") encourage people to take pledges for never picking up the phone while driving.

In general, the effect of these existing legislative, technological, and educational efforts seems to be inconclusive. Distracted driving affects drivers, their families, the workforce, and the society; therefore, everyone has a role to play and the responsibility to model positive behaviours (Gauld et al., 2019). More efforts are needed to develop long-lasting, effective, and sustainable programs and campaigns and the associated messaging and framing to really make a difference in changing behaviours, attitudes, and experiences (Arnold et al., 2019; Li et al., 2014).

The current study aimed to propose a new framework and include potential contributing factors that have not been explored. Specifically, we attempted to identify psychosocial factors that are relevant to social needs, habitual use, and self-control. We argued that one of the key psychological functions of cell phones or smartphones is its social function in the sense of fulfilling needs of belongingness regardless of physical place and building up social networks and connection to the world outside (Srivastava, 2005). Similarly, fear of missing out and social connectedness have been found to be related to phone use in general (Przybylski & Weinstein, 2013). The concept of belongingness and social needs are considered a driving force for individuals seeking social

connection via physical or virtual interactions with their social groups (Verduyn et al., 2017).

Another construct – habitual use – is related to attachment. Smartphones have been described as having characteristics of an attachment bond (Ainsworth, 1985): as digital companions (Carolus et al., 2019), digital pacifiers (Diefenbach & Borrmann, 2019; Melumad & Pham, 2017), forms of gratification (Fullwood et al., 2017), social entity (Carolus et al., 2019; Srivastava, 2005), reward mechanism (Aranda & Baig, 2018), and virtual friend (Fullwood et al., 2017). A third construct – self-control – is related to the fact that drivers are aware of the danger of problematic phone use while driving and the bans and regulations associated with it and yet they irrationally use the phones while driving. Drivers may pace their interactions with the phones or other distracting secondary tasks when they feel bored, traffic is light, road conditions are not challenging, or the importance of the phone call is high (Oviedo-Trespalacios et al., 2018; Wandtner et al., 2016). In other words, self-regulation plays a role in how and when drivers pick-up or put-down the phones and needs to strike a balance between the strength of an impulse and own ability to inhibit such desire (Kelley et al., 2015).

We hypothesized that social needs drive an individual’s habitual use of cell phones, and habitual use and self-control then lead to problematic phone use while driving (see Figure 1).

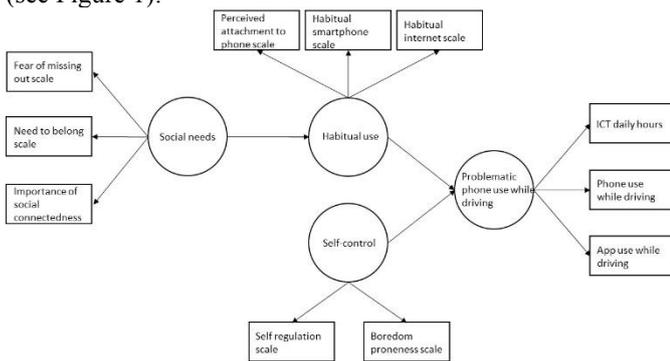


Fig. 1. Proposed framework

2. Method

An online, anonymous survey was used for this work. This study received the Institution Review Board approval from the first author’s university. This survey was administered via Qualtrics and posted on Mechanical Turk.

2.1 Participants

Individuals who held the status of an Amazon Mechanical Turk Master (Lovett et al., 2018) (workers who have demonstrated high performance over time and meet the performance requirements put forth by Mechanical Turk) were invited to participate. Individuals also had to be an adult and be residing in the U.S. Initially, 402 participants completed the survey and received the compensation of USD 7; however, 8 of them did not pass the attention check questions (e.g., answering 1978 when the survey asked for the current year) and were removed from the dataset. Thus, the final sample size was 394.

2.2 Procedures

Individuals who accepted the invitation to participate in this study were first directed to the consent page and must agree to the requirement of completing the entire survey. Once they indicated consent, they were asked to read the instructions as well as the definitions of the terminology used in the survey. The instructions also emphasized that they should answer the survey honestly and that there were no right or wrong answers. Survey questions were presented one at a time and participants could skip questions, although they were encouraged to answer all the questions. On average, participants took 16 min to complete the survey.

2.3 Key Measures

The survey included several previously validated psychosocial scales: The Need to Belong Scale (Baumeister & Leary, 1995), the Fear of Missing Out Scale (Przybylski et al., 2013), Perceived Attachment to Phone Scale (Weller et al., 2013), Habitual Smartphone/Internet Behaviour Scale (Limayem et al., 2003; van Deursen et al., 2015), the Self Regulation Scale (Diehl et al., 2006), and the Boredom Proneness Scale (Farmer & Sundberg, 1986; Vodanovich et al., 2005). The rest of the survey items were developed by the authors, including sources of news and information, importance of social connectedness, driving frequency, annual mileage, car accident involvement, phone use while driving frequency, mobile application use while driving, as well as demographic questions (participants’ age, gender, residence, state of residence, race and ethnicity, education, income, and employment status).

2.4 Analytical Strategy

Structural Equation Modelling approach is being used to examine the measurement models between the observed variables and the latent constructs as well as the relationships among the latent constructs. The proposed framework and the structural relationships among the variables are presented in Figure 1.

3. Results

3.1 Sample Characteristics

The sample consisted of 219 men and 175 women who lived in the U.S. and the ages ranged from 20 to 76, with the average being 40.89 (SD = 11.21) years. Participants came from 45 states (no data came from Alaska, Arkansas, North Dakota, South Dakota, and Vermont). Of the 394 responses, 190 chose “suburban,” 125 chose “urban,” 77 chose “rural,” and 2 chose “other” as their primary residence area. The majority of the participants identified their race and ethnicity as White (n = 307) (61 as Asian, 20 as Black, 12 as Hispanic/Latino/Spanish origin, 11 as American Indian/Alaska Native, 1 as Native Hawaiian/Other Pacific Islander, and 2 as Other). In terms of education level, 219 participants reported having a college degree, followed by having some college (n = 70), having a graduate degree (n = 56), having a high school diploma (n = 47), and having some high school education (n = 2). The annual household income item included five options: most of participants selected the 45-70 K (n = 115) and 25-45 K (n = 105) options, followed by the 70-110K option (n = 69), <25K option (n = 62),

and >110K option (n = 43). More than half of the participants reported having a full-time job (n = 258) (22 worked part-time, 82 were self-employed, 3 were a student, 27 were unemployed).

3.2 Structural Equation Models

Currently, the SEM is still being conducted and the results will be available when the full paper is due, should the abstract is accepted.

4. Discussion and Conclusions

Distracted driving as well as the precursors and consequences are complex and multifaceted behavioural, psychological, public health, and transportation safety issues. Cell phones afford many contemporary benefits but using them while driving a vehicle put the users and others in danger. As researchers consider distracted driving behaviours modifiable and preventable (Bingham, 2014; Brown et al., 2016), better understanding of the underlying psychological needs and motivations is critical in guiding the design of regulatory and educational efforts. There is a critical need to investigate the occurrence of interrelated distracting behaviours and how drivers change their communication preferences, as these behaviours may be motivated by some common needs. For example, the framing of messages, educational activities, and pledges can be better aligned with target users' psychological needs and motivations. "Your friends can wait" might be more comprehensively perceived than "It can wait" by acknowledging the social component and the significance to relationship maintenance.

References

- Ainsworth, M. D. (1985). Attachments across the life span. *Bulletin of the New York Academy of Medicine*, 61(9), 792–812.
- Aranda, J. H., & Baig, S. (2018). Toward "JOMO": The joy of missing out and the freedom of disconnecting. *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 19:1-19:8. <https://doi.org/10.1145/3229434.3229468>
- Arnold, L. S., Benson, A., Tefft, B. C., Barragan, D., Jin, L., Kolek, S., & Horrey, W. J. (2019). *Effectiveness of distracted driving countermeasures: A review of the literature (research brief)*. AAA Foundation for Traffic Safety. <https://aaaafoundation.org/effectiveness-of-distracted-driving-countermeasures-a-review-of-the-literature/>
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497–529. <https://doi.org/10.1037/0033-2909.117.3.497>
- Bingham, C. R. (2014). Driver distraction: A perennial but preventable public health threat to adolescents. *Journal of Adolescent Health*, 54(5, Supplement), S3–S5. <https://doi.org/10.1016/j.jadohealth.2014.02.015>
- Brown, D. T., George, D. A., Rickwood, P. D., & Frost, R. (2016). *Risk and regret: Identifying modifiable factors that predict drink driving and texting while driving among young adult drivers in the ACT region*.
- Caird, J. K., Johnston, K. A., Willness, C. R., Asbridge, M., & Steel, P. (2014). A meta-analysis of the effects of texting on driving. *Accident Analysis & Prevention*, 71, 311–318. <https://doi.org/10.1016/j.aap.2014.06.005>
- Carolus, A., Binder, J. F., Muench, R., Schmidt, C., Schneider, F., & Buglass, S. L. (2019). Smartphones as digital companions: Characterizing the relationship between users and their phones. *New Media & Society*, 21(4), 914–938. <https://doi.org/10.1177/1461444818817074>
- Creaser, J. I., Edwards, C. J., Morris, N. L., & Donath, M. (2015). Are cellular phone blocking applications effective for novice teen drivers? *Journal of Safety Research*, 54, 75.e29-78. <https://doi.org/10.1016/j.jsr.2015.06.014>
- Curtin, M. (2018, October 30). *Are you on your phone too much? The average person spends this many hours on it every day*. Inc.Com. <https://www.inc.com/melanie-curtin/are-you-on-your-phone-too-much-average-person-spends-this-many-hours-on-it-every-day.html>
- Diefenbach, S., & Borrmann, K. (2019). The smartphone as a pacifier and its consequences: Young adults' smartphone usage in moments of solitude and correlations to self-reflection. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 306:1-306:14. <https://doi.org/10.1145/3290605.3300536>
- Diehl, M., Semegon, A. B., & Schwarzer, R. (2006). Assessing attention control in goal pursuit: A component of dispositional self-regulation. *Journal of Personality Assessment*, 86(3), 306–317. https://doi.org/10.1207/s15327752jpa8603_06
- Farmer, R., & Sundberg, N. D. (1986). Boredom proneness—The development and correlates of a new scale. *Journal of Personality Assessment*, 50(1), 4–17. https://doi.org/10.1207/s15327752jpa5001_2
- Fullwood, C., Quinn, S., Kaye, L. K., & Redding, C. (2017). My virtual friend: A qualitative analysis of the attitudes and experiences of Smartphone users: Implications for Smartphone attachment. *Computers in Human Behavior*, 75, 347–355. <https://doi.org/10.1016/j.chb.2017.05.029>
- Gauld, C. S., Lewis, I. M., White, K. M., Fleiter, J. J., & Watson, B. (2019). Public education messages aimed at smartphone use among young drivers: A mixed methods exploration of their effectiveness. *Transportation Research Part F: Traffic Psychology and Behaviour*, 60, 311–326. <https://doi.org/10.1016/j.trf.2018.10.027>
- Highway Loss Data Institute. (2009). Hand-held cellphone laws and collision claim frequencies. *Highway Loss Data Institute Bulletin*, 26(17).
- Kashevnik, A., Shchedrin, R., Kaiser, C., & Stocker, A. (2021). Driver Distraction Detection Methods: A Literature Review and Framework. *IEEE Access*, 9, 60063–60076. <https://doi.org/10.1109/ACCESS.2021.3073599>

- Kelley, W. M., Wagner, D. D., & Heatherton, T. F. (2015). In search of a human self-regulation system. *Annual Review of Neuroscience*, 38(1), 389–411. <https://doi.org/10.1146/annurev-neuro-071013-014243>
- Leipnitz, J., Gross, A., Dostert, J., & Baumgarten, T. (2022). Evaluation of a dynamic blocking concept to mitigate driver distraction: Three simulator studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, 84, 262–276. <https://doi.org/10.1016/j.trf.2021.12.004>
- Li, W., Gkritza, K., & Albrecht, C. (2014). The culture of distracted driving: Evidence from a public opinion survey in Iowa. *Transportation Research Part F: Traffic Psychology and Behaviour*, 26, 337–347. <https://doi.org/10.1016/j.trf.2014.01.002>
- Lim, S. H., & Chi, J. (2013). Are cell phone laws in the U.S. effective in reducing fatal crashes involving young drivers? *Transport Policy*, 27, 158–163. <https://doi.org/10.1016/j.tranpol.2013.01.011>
- Limayem, M., Hirt, S., & Cheung, C. (2003). Habit in the context of IS continuance: Theory extension and scale development. *ECIS 2003 Proceedings*. <https://aisel.aisnet.org/ecis2003/90>
- Lovett, M., Bajaba, S., Lovett, M., & Simmering, M. J. (2018). Data quality from crowdsourced surveys: A mixed method inquiry into perceptions of Amazon's Mechanical Turk Masters. *Applied Psychology*, 67(2), 339–366. <https://doi.org/10.1111/apps.12124>
- McCartt, A. T., & Geary, L. L. (2004). Longer term effects of New York State's law on drivers' handheld cell phone use. *Injury Prevention*, 10(1), 11–15. <https://doi.org/10.1136/ip.2003.003731>
- McDonald, C. C., Ward, K., Huang, Y., Wiebe, D. J., & Delgado, M. K. (2019). Novel smartphone-based measures of cell phone use while driving in a sample of newly licensed adolescent drivers. *Health Education & Behavior*, 46(1), 10–14. <https://doi.org/10.1177/1090198118788612>
- Melumad, S., & Pham, M. T. (2017). *Understanding the psychology of smartphone usage: The adult pacifier hypothesis* (NA-Advances in Consumer Research Volume 45; Eds. Ayelet Gneezy, Vlaslas Griskevicius, and Patti Williams, pp. 25–30). Association for Consumer Research.
- Oviedo-Trespalacios, O., Haque, M. M., King, M., & Demmel, S. (2018). Driving behaviour while self-regulating mobile phone interactions: A human-machine system approach. *Accident Analysis & Prevention*, 118, 253–262. <https://doi.org/10.1016/j.aap.2018.03.020>
- Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V. (2013). Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior*, 29(4), 1841–1848. <https://doi.org/10.1016/j.chb.2013.02.014>
- Przybylski, A. K., & Weinstein, N. (2013). Can you connect with me now? How the presence of mobile communication technology influences face-to-face conversation quality. *Journal of Social and Personal Relationships*, 30(3), 237–246. <https://doi.org/10.1177/0265407512453827>
- Rocco, L., & Sampaio, B. (2016). Are handheld cell phone and texting bans really effective in reducing fatalities? *Empirical Economics*, 51(2), 853–876. <https://doi.org/10.1007/s00181-015-1018-8>
- Srivastava, L. (2005). Mobile phones and the evolution of social behaviour. *Behaviour & Information Technology*, 24(2), 111–129. <https://doi.org/10.1080/01449290512331321910>
- Stadd, A. (2013). *79% of people 18-44 have their smartphones with them 22 hours a day*. Adweek.Com. <https://www.adweek.com/digital/smartphones/>
- Strayer, D. L., & Fisher, D. L. (2016). SPIDER: A framework for understanding driver distraction. *Human Factors*, 58(1), 5–12. <https://doi.org/10.1177/0018720815619074>
- van Deursen, A. J. A. M., Bolle, C. L., Hegner, S. M., & Kommers, P. A. M. (2015). Modeling habitual and addictive smartphone behavior: The role of smartphone usage types, emotional intelligence, social stress, self-regulation, age, and gender. *Computers in Human Behavior*, 45, 411–420. <https://doi.org/10.1016/j.chb.2014.12.039>
- Verduyn, P., Ybarra, O., Résibois, M., Jonides, J., & Kross, E. (2017). Do social network sites enhance or undermine subjective well-being? A critical review. *Social Issues and Policy Review*, 11(1), 274–302. <https://doi.org/10.1111/sipr.12033>
- Vodanovich, S. J., Wallace, J. C., & Kass, S. J. (2005). A confirmatory approach to the factor structure of the Boredom Proneness Scale: Evidence for a two-factor short form. *Journal of Personality Assessment*, 85(3), 295–303. https://doi.org/10.1207/s15327752jpa8503_05
- Wandtner, B., Schumacher, M., & Schmidt, E. A. (2016). The role of self-regulation in the context of driver distraction: A simulator study. *Traffic Injury Prevention*, 17(5), 472–479. <https://doi.org/10.1080/15389588.2015.1102231>
- Weller, J. A., Shackelford, C., Dieckmann, N., & Slovic, P. (2013). Possession attachment predicts cell phone use while driving. *Health Psychology*, 32(4), 379–387. <https://doi.org/10.1037/a0029265>
- World Bank. (2019). *Mobile cellular subscriptions (per 100 people)*. <https://data.worldbank.org/indicator/IT.CEL.SETS.P2>
- World Health Organization, & National Highway Traffic Safety Administration. (2011). *Mobile phone use: A growing problem of driver distraction*. Geneva: World Health Organization. <http://apps.who.int/iris/handle/10665/44494>

TXT N DRV: A Systematic Review of the Effect of Texting while Driving on Driver Performance

Burcu Arslan^{1*}

¹Department of Psychology, Middle East Technical University, Ankara, TURKEY

[*barslan.psy@gmail.com](mailto:barslan.psy@gmail.com)

Abstract: A growing body of research demonstrates the effect of texting while driving (TWD) on driver performance. It is shown that TWD is responsible for poor driver performance and increased accident probability. The current review aims to explore the articles examining the effect of TWD on driver performance, and to create a holistic picture of this relationship. Twenty three articles met the selection criteria of this review. The articles are classified into two categories. At the first category, the articles examining the effect of only TWD on driver performance is included. At the second category, the articles comparing the effect of handheld and hands-free TWD on driver performance is analyzed. Driver performance was measured based on six indicators, which are accident probability, lateral vehicle control, longitudinal vehicle control, reaction time, visual scanning behavior and workload of drivers. The review revealed that the TWD had a great effect on driver's accident probability, lateral vehicle control, longitudinal vehicle control, reaction time, visual scanning behavior and cognitive workload. Also, studies comparing the effect of handheld and hands-free TWD showed that the handheld TWD has the most degrading effect on driver performance.

1. Introduction

Of all driver distractions, TWD is considered the most dangerous driver distraction (Caird et al., 2014). Drivers have to move their eyes from road to phone, comprehend the content in the message and typed to reply, and thus may lose control over the steering wheel at hand. This process highlights that TWD results in at least three distraction types; cognitive, visual, and physical distraction. Notwithstanding its highly dangerous nature, TWD is becoming more popular, especially among young drivers. The study conducted with high school students in the USA revealed that 38% of 101,397 participants reported TWD at least once (Li et al., 2018). Another study surveying US drivers of all ages showed that 60% of 1211 drivers reported TWD, whereas the highest number of cell phone use while driving belonged to young drivers (Gliklich, et al., 2016). Haste (2005) reported critically high TWD among UK drivers, which was about nine out of ten drivers (89%). Hill, et al. (2019) stated that one-third of 220 participants reported engaging in TWD

daily in Ukraine. Data from South Africa also presented a similar pattern. 60% of drivers recruited in the study accepted that they often engage in TWD (Oyedemi and Kgasago, 2017). These studies showed that these alarming rates of TWD are not regional but a global issue. Thus, understanding its impact on driver performance has attracted researchers for more than a decade.

The studies regarding its impact on driver performance clearly showed the magnitude of this threat. Numerous studies utilizing driving simulator or instrumented car was published and addressed similar problems such as decrements in control over the vehicle (Yannis et al., 2016; McKeever, Schultheis et al., 2013), increased crash risk (Lansdown, 2019), degrading visual scanning performance (Rudin-Brown et al., 2013) and increase in drivers' workload (Young, et al., 2014) and reaction time (Lyngsie, et al., 2013).

There are many studies reviewing driver distraction or mobile phone use while driving and their impact on driver performance (Young et al., 2007; Oviedo-Trespalacios et al., 2016).

However, to the best of our knowledge, there was a lack of comprehensive review study focusing on the impact of TWD on driver performance, despite the popularity and seriousness of this issue.

The current study aimed to review studies investigating the texting-while-driving (TWD) behavior and driver performance. Thus, the studies measuring drivers' performance under the influence of TWD texting will be presented and their findings will be synthesized in following sections.

2. Method

The literature related to texting while driving (TWD) and driver performance was examined. Peer-reviewed articles were obtained via Scopus database (www.scopus.com). Firstly, "texting while driving" is used as an only keyword since it is the most common term referring to the texting behavior in driving context. Secondly, "texting" is combined with either "driving" or "driver performance" in order not to miss any related documents. Additionally, the search was done by selecting "title, abstract, keyword" option. The duration of documents were set as "all years" to present while document type was selected as "article" and "article in press". Finally, the search was resulted in 487 article in total. The several articles were excluded from study based on which i) the language of the document was not English ii) they were duplicated iii) full-text and abstracts could not be reached iv) the focus of study was not on texting but was on mobile phone use in general or distracted driving v) "texting while driving" was the dependent variable. After the exclusion process, twenty three articles remained to examine.

3. Results & Conclusions

One of the most salient effects of TWD was on accident probability. Drews et al.'s (2009) study showed that almost all accidents occurred while drivers were texting. Bendak (2015) and Lansdown (2009) also showed that TWD had increased the accident number more than four times and five times, compared to driving without distraction. Additionally, the

TWD related accident numbers were reported even higher on relatively riskier road environments. For example, the accident risk was higher on urban roads compared to rural roads (Yannis et al., 2014), on highways compared to town roads (Bendak, 2015), and in heavy traffic compared to moderate traffic situations (Yannis et al., 2016).

Furthermore, the lateral vehicle control performance was also degraded by TWD and contributed to the crash risks (Lyngsie et al., 2013; McKeever et al., 2013). The majority of the studies showed that TWD significantly affected the lane management performance. The drivers did not seem to be aware of the car's position within lanes or the steering wheel's position at hand (Rudin-Brown et al., 2013). In parallel to accident probability, the lane deviation due to TWD was higher in high-risk road conditions, such as in tunnels (Rudin-Brown et al., 2013), and on city roads with many stimuli (Lyngsie et al., 2013). Another important finding was that the impairments in lateral vehicle control continued for a little longer after TWD ended (Thapa et al., 2015). That is, drivers needed more time to reattend the road conditions after they unhandled the phone

The impairments in visual scanning behavior underlied how distracted driver's accident involvement was paved with TWD. Related studies showed that drivers' eyes off the road significantly contributed to poor performance and crash risk (Lansdown, 2001). Drivers kept their eyes off the road for 31% of driving time on highways (Bendak, 2015), or that number of glances off the road was 30-50 times higher while TWD compared to undistracted driving. Thapa et al. (2015) also suggested that the continued degradations in driver performance after the texting sessions might be due to the residual effect of fixing eyes back to the road from the phone.

It was not surprising since texting requires drivers to divert their attention to the phone instead of the road and make mental effort to comprehend the text's content and reply accordingly. It should be noted that the drivers' reaction time was measured by objective

indicators like brake response to front cars' brake lights, whereas workload was measured by a self-report method. It was assumed that the drivers recognized how effortful the driving safely while texting was.

The effect of TWD on longitudinal vehicle control was not considered degrading the driver performance, if statistically speaking. The majority of studies showed that TWD engagement decreased vehicles' speed and increased the following distance and headway time/distance (Choudhary and Velaga, 2017; Yannis et al., 2014, etc.). These findings were attributed to the drivers' attempt to compensate for the risks of TWD. Caird et al. (2014) suggested that an increase in the following distance could be the attempt to create a safety buffer to reduce the crash risk. Similarly, Rudin-Brown et al. (2013) stated that, the speed reduction was higher in TWD conditions in high-risk road environments such as tunnels, than in low-risk road environments. It showed that the drivers' compensation effort increased along with the increased risks on the road.

Four studies reviewed here were conducted to examine whether texting via a speech-based interface and an in-vehicle system being safer than texting with a handheld phone. In general, there was a evidence that handheld texting was more degrading than texting with other modalities in all driver performance indicators (Terken et al., 2011; Chen et al., 2020; He et al., 2014; Owens et al., 2011). Chen et al. (2020) suggested that handheld texting was riskier and required more effort to compensate for the accident probability. However, these results did not imply that hand-free texting contributed to safe driving. The accident probability, drivers' mental workload, lane deviations, and reaction times were still increased in other texting conditions, compared to undistracted driving (Terken et al., 2011; Chen et al., 2020; He et al., 2014; Owens et al., 2011).

Concerning the findings above, it is clear that TWD is a highly threatening risk for the drivers' safety. Its effect can be observed on almost every performance indicator directly or indirectly thorough another performance

indicator. For example, the significant effect of TWD on the lateral control can be related to the duration of drivers' eyes off the road. Hence, while planning countermeasures for TWD, it should be noted that targeting one or two performance indicators could be ineffective.

Despite knowing the risks of TWD, the newer and presumably safer ways to keep drivers "in touch" is being developed every day. Some studies show that, in some cases, the effect of TWD can be reduced by using new methods of hands-free texting, such as speech-based texting or texting via an in-vehicle system. It should be kept in mind that these methods are not risk-free; indeed, they carry a considerable risk of accidents (He et al., 2014) and become a burden for road safety. Thus, instead of technological developments regarding new ways of texting within vehicles, road safety researchers should emphasize the TWD banning legislation, enforcements, and interventions.

References

- Bendak, S. (2015). Objective assessment of the effects of texting while driving: a simulator study. *International journal of injury control and safety promotion*, 22(4), 387-392.
- Caird, J. K., Johnston, K. A., Willness, C. R., Asbridge, M., & Steel, P. (2014). A meta-analysis of the effects of texting on driving. *Accident Analysis & Prevention*, 71, 311-318.
- Chen, Y., Fu, R., Xu, Q., & Yuan, W. (2020). Mobile Phone Use in a Car-Following Situation: Impact on Time Headway and Effectiveness of Driver's Rear-End Risk Compensation Behavior via a Driving Simulator Study. *International journal of environmental research and public health*, 17(4), 1328.
- Choudhary, P., & Velaga, N. R. (2017). Mobile phone use during driving: Effects on speed and effectiveness of driver compensatory behaviour. *Accident Analysis & Prevention*, 106, 370-378.
- Draws, F. A., Yazdani, H., Godfrey, C. N., Cooper, J. M., & Strayer, D. L. (2009). Text messaging

- during simulated driving. *Human factors*, 51(5), 762-770.
- Gliklich, E., Guo, R., & Bergmark, R. W. (2016). Texting while driving: A study of 1211 US adults with the Distracted Driving Survey. *Preventive medicine reports*, 4, 486-489.
- Haste, H. (2005). Joined-up texting: Mobile phones and young people. *Young Consumers: Insight and Ideas for Responsible Marketers*, 6(3), 56-67.
- He, J., Chaparro, A., Nguyen, B., Burge, R. J., Crandall, J., Chaparro, B., ... & Cao, S. (2014). Texting while driving: Is speech-based text entry less risky than handheld text entry?. *Accident Analysis & Prevention*, 72, 287-295.
- Hill, T., Sullman, M. J. M., & Stephens, A. N. (2019). Mobile phone involvement, beliefs, and texting while driving in Ukraine. *Accident Analysis & Prevention*, 125, 124-131.
- Lansdown, T. C. (2019). The temptation to text when driving—Many young drivers just can't resist. *Transportation research part F: traffic psychology and behaviour*, 65, 79-88.
- Li, P., Markkula, G., Li, Y., & Merat, N. (2018). Is improved lane keeping during cognitive load caused by increased physical arousal or gaze concentration toward the road center?. *Accident Analysis & Prevention*, 117, 65-74.
- Lyngsie, K., Pedersen, M. S., Stage, J., & Vestergaard, K. F. (2013, September). Don't text while driving: the effect of smartphone text messaging on road safety during simulated driving. In *IFIP Conference on Human-Computer Interaction* (pp. 546-563). Springer, Berlin, Heidelberg.
- McKeever, J. D., Schultheis, M. T., Padmanaban, V., & Blasco, A. (2013). Driver performance while texting: even a little is too much. *Traffic injury prevention*, 14(2), 132-137.
- Oviedo-Trespalacios, O., Haque, M. M., King, M., & Washington, S. (2016). Understanding the impacts of mobile phone distraction on driving performance: A systematic review. *Transportation research part C: emerging technologies*, 72, 360-380.
- Owens, J. M., McLaughlin, S. B., & Sudweeks, J. (2011). Driver performance while text messaging using handheld and in-vehicle systems. *Accident Analysis & Prevention*, 43(3), 939-947.
- Oyedemi, T., & Kgasago, T. J. (2017). Always-Available Communication and Technological Distractions: Technology Use, Texting and Driving. *Communicatio*, 43(3-4), 36-53.
- Rudin-Brown, C. M., Young, K. L., Patten, C., Lenné, M. G., & Ceci, R. (2013). Driver distraction in an unusual environment: Effects of text-messaging in tunnels. *Accident Analysis & Prevention*, 50, 122-129.
- Terken, J., Visser, H. J., & Tokmakoff, A. (2011, November). Effects of speech-based vs handheld e-mailing and texting on driving performance and experience. In *Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 21-24).
- Thapa, R., Codjoe, J., Ishak, S., & McCarter, K. S. (2015). Post and during event effect of cell phone talking and texting on driving performance—A driving simulator study. *Traffic injury prevention*, 16(5), 461-467.
- Yannis, G., Laiou, A., Papantoniou, P., & Christoforou, C. (2014). Impact of texting on young drivers' behavior and safety on urban and rural roads through a simulation experiment. *Journal of safety research*, 49, 25-e1.
- Yannis, G., Laiou, A., Papantoniou, P., & Gkartzonikas, C. (2016). Simulation of texting impact on young drivers' behavior and safety on motorways. *Transportation research part F: traffic psychology and behaviour*, 41, 10-18.
- Young, K. L., Rudin-Brown, C. M., Patten, C., Ceci, R., & Lenné, M. G. (2014). Effects of phone type on driving and eye glance behaviour while text-messaging. *Safety science*, 68, 47-54.
- Young, K., Regan, M., & Hammer, M. (2007). Driver distraction: A review of the literature. *Distracted driving, 2007*, 379-405

Distracted, unfocused and risky driving in adolescents: Associations with sleep issues, inattention and sluggish cognitive tempo symptoms

Vanessa Beanland^{1*}, Catherine Ritchie¹, Barbara C. Galland², Elizabeth A. Schaughency¹

¹ Department of Psychology, University of Otago, Dunedin, New Zealand

² Department of Women's and Children's Health, University of Otago, Dunedin, New Zealand

* vanessa.beanland@otago.ac.nz

Abstract: Road crashes are a leading cause of death and disability in adolescence. Various factors can contribute to these crashes including overt risk-taking but also distracted, inattentive, and unfocused driving. Adolescent drivers are likely to be more susceptible to these issues if they experience sleep issues or have symptoms associated with inattention (e.g., ADHD) or sluggish cognitive tempo (SCT). In this study we surveyed 365 teenage drivers aged 16-18 years living in New Zealand. The survey included measures of aberrant driving, self-reported sleep habits and issues, inattention and SCT. Smartphone use while driving was significantly associated with sleep issues, inattention and SCT symptoms among teens who drive unsupervised, but not those who only drive under supervision (i.e., Learners). Unfocused driving showed small correlations with daytime sleepiness, inattention and SCT symptoms among all teen drivers. Overt risk-taking while driving showed no correlations with sleep issues, and only small correlations with SCT symptoms (for all teen drivers) and inattention (for unsupervised drivers only). The results highlight that sleep issues, inattention symptoms and SCT symptoms are key risk factors for both distracted and inattentive driving among young drivers, especially those driving without supervision.

1. Introduction

Road crashes are a leading cause of death and disability in adolescence (Toroyan & Peden, 2007). Risky driving includes not only overt risk-taking while driving (i.e., speeding, aggressive driving, and driving after substance use), but also inattentive, distracted, and unfocused driving (Nadaraja et al., 1997). Risky and inattentive driving, as well as insufficient sleep and daytime sleepiness, are common contributory factors to road crashes (Martiniuk et al., 2013; Shope et al., 2008).

Adolescents with conditions or characteristics affecting attention, such as attention deficit hyperactivity disorder (ADHD) or sluggish cognitive tempo (SCT) symptoms, may be more susceptible to some forms of risky driving, especially unfocused driving and driver distraction. SCT often overlaps with but is psychometrically distinct from the ADHD-I inattentive subtype; it is characterized by symptoms of daydreaming, daytime sleepiness, slowed thinking, and lethargy (Müller et al., 2014). SCT symptoms have been associated with functional impairments, including executive dysfunction, emotional dysregulation, and daytime sleepiness (Flannery et al., 2017). These issues are likely to be exacerbated by sleep issues; ADHD and SCT symptoms may act as mediators in the relation between poor sleep and risky driving (Jerome et al., 2006; Watling & Watling, 2021).

The current online survey study was designed to explore associations sleep issues, inattention and SCT symptoms have with aberrant driving in teenaged drivers, especially distracted, unfocused and risky driving.

The research was conducted in New Zealand, where teenagers can obtain a Learner licence at 16 years, a Restricted licence (which permits unsupervised driving under most conditions) at 16.5 years, and a full licence at either 17.5 years (if they complete an approved driver training course) or 18 years (default minimum age).

2. Method

2.1 Participants

The sample included 365 teenager drivers aged 16-18 years ($M = 17$ years 3 months, $SD = 8.3$ months) who lived in New Zealand. This included 159 teens with a Learner licence who only drove under supervision, 148 with a Restricted licence, 30 with a full licence, and 28 who reported driving unsupervised despite having no licence or only a Learner licence.

There were 196 females (54%), 159 males (44%), and 10 individuals who identified as gender diverse or did not specify gender. Participants self-reported their ethnicities as 84.7% New Zealand European, 11.0% Māori, 3.0% Indian, 2.2% Samoan, 1.6% Chinese, and 11.5% other (multiple ethnicity selections were permitted).

Ethical aspects of the research were approved by the University of Otago Human Ethics Committee (21/031). Participants were recruited via online ads on Facebook and Instagram and through targeted emails to New Zealand secondary schools, with data collection occurring from May to August 2021.

2.2 Measures

Participants completed measures to assess self-reported sleep duration, sleep issues and daytime sleepiness, SCT symptoms, general inattention, and driving behaviour including risky, distracted, and unfocused driving.

To assess sleep issues, participants were asked if they had difficulty initiating sleep, staying asleep, and/or returning to sleep following waking at least three nights a week over three months or longer. These questions map onto DSM-5 criteria for risk of insomnia (APA, 2013).

The Epworth Sleepiness Scale for Children and Adolescents (ESS-CHAD; Janssen et al., 2017) was used to

assess daytime sleepiness. Participants rated their likelihood of falling asleep in eight situations on a 4-point scale from 0 (*would never fall asleep*) to 3 (*high chance of falling asleep*). Scores are summed across items. Possible scores range from 0 to 24 with higher scores indicating increased daytime sleepiness.

SCT symptoms were measured using the Child Concentration Inventory 2nd edition (CCI-2; Becker, 2015), which requires participants to rate how frequently they experience symptoms (e.g., “My mind feels like it is in a fog”) on a 4-point scale from 0 (*never*) to 3 (*always*). The CCI-2 contains 16 items, but previous research has shown three items do not show discriminant and convergent validity (Becker et al., 2020) so only 13 items were used to calculate CCI-2 summary scores. Possible scores range from 0 to 39 with higher scores indicating more SCT symptoms.

Attentional difficulties were assessed with a 9-item self-report adaptation of the inattention items employed by Beebe et al. (2008) from the Vanderbilt Assessment Scale (Wolraich et al., 2003). Attention items (e.g., “Had difficulty keeping attention to what needed to be done”) were rated on a 4-point scale from 0 (*never*) to 3 (*very often*). Possible scores range from 0 to 27 with higher scores indicating more inattention symptoms.

Distracted and unfocused driving were assessed using the Behaviour of Young Novice Drivers Scale (BYNDS; Scott-Parker et al., 2010; Scott-Parker & Proffitt, 2015). Specifically, the Unfocused Driving (BYNDS-UD; 13-items) and Smartphone Use (BYNDS-SMP; 7-items) subscales from an updated version of the BYNDS (Jannusch et al., 2020) were used for analyses. Participants rated driving behaviours (e.g., “You misjudge the speed of an oncoming vehicle”, “You read messages on your smartphone”) on a 5-point scale from 1 (*never*) to 5 (*nearly all the time*).

Risky driving was assessed using the risky driving subscale of the Dula Dangerous Driving Index (DDDI-RD; Dula & Ballard, 2003). This included 12 items (e.g., “I will drive if I am only mildly intoxicated or buzzed”) rated on a 5-point scale from 1 (*never*) to 5 (*always*). For comparability DDDI and BYNDS scores were averaged across items to derive subscale scores between 1 and 5, with higher scores representing more frequent aberrant driving.

2.3 Procedure

Participants completed an anonymous online survey through REDCap (Research Electronic Data Capture; Harris

et al., 2009), a secure online platform used to create and manage online databases and surveys.

After indicating consent participants completed the self-report measures in 2.2, then demographic questions. Participants were then taken to a separate survey where they could enter a prize draw. Finally, participants were presented with links to websites on safe driving.

3. Results

3.1 Descriptive statistics

All scales showed good reliability (see Table 1) and the full possible range was observed (i.e., there were participants at both extremes of all scales).

Over half the participants (55.6%) were classified as having sleep issues, in that they reported having initiating sleep, remaining asleep and/or returning to sleep.

Table 1 Descriptive statistics for each measure

| Scale | <i>M</i> (<i>SD</i>) | Range | Cronbach's alpha |
|--------------------------|------------------------|-------|------------------|
| ESS-CHAD | 7.1 (4.2) | 0–24 | .76 |
| CCI-2 | 16.8 (9.0) | 0–39 | .95 |
| Inattention (Vanderbilt) | 8.5 (5.7) | 0–27 | .90 |
| BYNDS-UD | 1.8 (0.5) | 1–5 | .82 |
| BYNDS-SMP | 1.8 (0.8) | 1–5 | .91 |
| DDDI-RD | 1.3 (0.5) | 1–5 | .91 |

3.2 Correlations

Because all three driving-related subscales were positively skewed, and insomnia symptoms was a binary category, Spearman's rho was used to correlate driving behaviour with sleep, inattention and SCT symptoms. Because driving behaviour differs between supervised learners and unsupervised drivers, the sample was split into two subgroups.

As shown in Table 2, smartphone use while driving was significantly positively correlated with self-reported insomnia, daytime sleepiness, inattention and SCT symptoms in unsupervised but not supervised drivers.

In contrast, unfocused driving showed similar patterns for unsupervised and supervised drivers: small correlations with daytime sleepiness and small to moderate correlations with inattention and SCT symptoms, although the correlations were larger for unsupervised drivers.

Table 2 Spearman's rho correlations [and 95% confidence intervals] between measures

| | Insomnia | ESS-CHAD | CCI-2 | Inattention |
|-------------------------------------|------------------|-----------------|-----------------|-----------------|
| <i>Smartphone use (BYNDS-SMP)</i> | | | | |
| Restricted/Full/unsupervised | .18 [.04, .31] | .23 [.09, .36] | .21 [.07, .34] | .20 [.06, .33] |
| Supervised learner drivers | -.01 [-.17, .16] | .08 [-.08, .24] | .06 [-.10, .22] | .12 [-.05, .27] |
| <i>Unfocused driving (BYNDS-UD)</i> | | | | |
| Restricted/Full/unsupervised | .10 [-.04, .24] | .22 [.08, .35] | .39 [.26, .50] | .37 [.24, .49] |
| Supervised learner drivers | .10 [-.07, .26] | .19 [.03, .34] | .24 [.08, .39] | .31 [.15, .45] |
| <i>Risky driving (DDDI-RD)</i> | | | | |
| Restricted/Full/unsupervised | .02 [-.12, .16] | .05 [-.09, .19] | .16 [.02, .30] | .20 [.06, .33] |
| Supervised learner drivers | .03 [-.13, .19] | .10 [-.06, .26] | .06 [-.10, .22] | .27 [.11, .41] |

Risky driving was not correlated with insomnia or daytime sleepiness but did show small correlations with inattention for both groups and a small correlation with SCT for unsupervised drivers only.

4. Discussion

The current results indicate that distracted and unfocused driving in teenagers is associated with SCT and inattention symptoms and daytime sleepiness, especially for unsupervised drivers. Further analyses (in the full paper) will explore potential moderators and mediators of these relationships. The fact that these symptoms were not significantly correlated with smartphone use for supervised learners suggests that supervisors may attenuate some, but not all, undesirable driving behaviours.

Notably, most of the sample were classified as having symptoms of insomnia. This high percentage is concerning but consistent with previous research in New Zealand adolescents (Galland et al., 2020). There are biological and social explanations for why adolescents may experience inadequate sleep.

The construct of unfocused driving measured in the current study represents primarily unintentional errors and lapses, such as misjudging driving situations or diverting attention to irrelevant stimuli like music or roadside advertisements. This subscale showed the strongest correlations observed in the study, with inattention and SCT symptoms for unsupervised drivers. In contrast, the more volitional risky driving behaviours measured by the DDDI-RD subscale showed very few significant associations, with only small associations mainly for inattention. This shows that adolescents experiencing inattention and SCT symptoms are not necessarily at risk for dangerous driving in general, but for specific forms of aberrant driving that are directly related to their cognitive symptoms and characteristics.

5. Conclusions

Sleep issues, inattention symptoms and SCT symptoms are key risk factors for both distracted and inattentive driving among young drivers, especially those driving without supervision.

References

- American Psychiatric Association (APA). (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5)*. Washington, DC: American Psychiatric Publishing. <https://doi.org/10.1176/appi.books.9780890425596>
- Becker, S. P. (2015). *Child concentration inventory, second edition (CCI-2)*. Cincinnati, OH: Author.
- Becker, S. P., Burns, G. L., Smith, Z. R., & Langberg, J. M. (2020). Sluggish cognitive tempo in adolescents with and without ADHD: Differentiation from adolescent-reported ADHD inattention and unique associations with internalizing domains. *Journal of Abnormal Child Psychology, 48*(3), 391–406. <https://doi.org/10.1007/s10802-019-00603-9>
- Beebe, D. W., Fallone, G., Godiwala, N., Flanigan, M., Martin, D., Schaffner, L., & Amin, R. (2008). Feasibility and behavioral effects of an at-home multi-night sleep restriction protocol for adolescents. *Journal of Child Psychology and Psychiatry, 49*(9), 915–923. <https://doi.org/10.1111/j.1469-7610.2008.01885.x>
- Dula, C. S., & Ballard, M. E. (2003). Development and evaluation of a measure of dangerous, aggressive, negative emotional, and risky driving. *Journal of Applied Social Psychology, 33*(2), 263–282. <https://doi.org/10.1111/j.1559-1816.2003.tb01896.x>
- Flannery, A. J., Luebke, A. M., & Becker, S. P. (2017). Sluggish cognitive tempo is associated with poorer study skills, more executive functioning deficits, and greater impairment in college students. *Journal of Clinical Psychology, 73*(9), 1091–1113. <https://doi.org/10.1002/jclp.22406>
- Galland, B. C., de Wilde, T., Taylor, R. W., & Smith, C. (2020). Sleep and pre-bedtime activities in New Zealand adolescents: Differences by ethnicity. *Sleep Health, 6*(1), 23–31. <https://doi.org/10.1016/j.sleh.2019.09.002>
- Jannusch, T., Völler, M., Murphy, F., & Mullins, M. (2020). A new version of the Behaviour of Young Novice Drivers Scale (BYNDS). Insights from a randomised sample of 700 German young novice drivers. *Accident Analysis and Prevention, 145*, 105622. <https://doi.org/10.1016/j.aap.2020.105622>
- Janssen, K. C., Phillipson, S., O'Connor, J., & Johns, M. W. (2017). Validation of the Epworth Sleepiness Scale for Children and Adolescents using Rasch analysis. *Sleep Medicine, 33*, 30–35. <https://doi.org/10.1016/j.sleep.2017.01.014>
- Jerome, L., Segal, A., & Habinski, L. (2006). What we know about ADHD and driving risk: A literature review, meta-analysis and critique. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 15*(3), 105–125. <https://doi.org/10.1007/s11920-006-0045-8>
- Martiniuk, A. L. C., Senserrick, T., Lo, S., Williamson, A., Du, W., Grunstein, R. R., Woodward, M., Glozier, N., Stevenson, M., Norton, R., & Ivers, R. Q. (2013). Sleep-deprived young drivers and the risk for crash: The DRIVE Prospective Cohort Study. *JAMA Pediatrics, 167*(7), 1–8. <https://doi.org/10.1001/jamapediatrics.2013.1429>
- Müller, K., Tucha, L., Koerts, J., Groen, Y., Lange, K. W., & Tucha, O. (2014). Sluggish cognitive tempo and its neurocognitive, social and emotive correlates: A systematic review of the current literature. *Journal of Molecular Psychiatry, 2*(1), 5. <https://doi.org/10.1186/2049-9256-2-5>
- Nada-Raja, S., Langley, J. D., McGee, R. O. B., Williams, S. M., Begg, D. J., & Reeder, A. I. (1997). Inattentive and hyperactive behaviors and driving offenses in adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*(4), 515–522. <https://doi.org/10.1097/00004583-199704000-00014>
- Scott-Parker, B., & Proffitt, C. (2015). Validation of the Behaviour of Young Novice Drivers Scale (BYNDS) in a New Zealand young driver population. *Accident Analysis and Prevention, 77*, 62–71. <https://doi.org/10.1016/j.aap.2015.01.019>
- Scott-Parker, B., Watson, B., & King, M. (2010). The risky behaviour of young drivers: Developing a measurement tool. In Proceedings of the 20th

- Canadian Multidisciplinary Road Safety Conference (pp. 1–16). Canadian Association of Road Safety Professionals.
<https://eprints.qut.edu.au/34418/1/c34418.pdf>
- Shope, J. T., & Bingham, C. R. (2008). Teen driving: Motor-vehicle crashes and factors that contribute. *American Journal of Preventive Medicine, 35*(3), S261–S271.
<https://doi.org/10.1016/j.amepre.2008.06.022>
- Toroyan, T. & Peden, M. (Eds.) (2007). *Youth and Road Safety*. World Health Organization. Retrieved from www.who.int/violence_injury_prevention/publications
- Watling, C. N., & Watling, H. A. (2021). Younger drivers executive functioning and the relationship with experiencing signs of sleepiness. *Transportation Research. Part F: Traffic Psychology and Behaviour, 80*, 359–367.
<https://doi.org/10.1016/j.trf.2021.05.003>
- Wolraich, M. L., Lambert, W., Doffing, M. A., Bickman, L., Simmons, T., & Worley, K. (2003). Psychometric properties of the Vanderbilt ADHD Diagnostic Parent Rating Scale in a referred population. *Journal of Pediatric Psychology, 28*(8), 559–567.
<https://doi.org/10.1093/jpepsy/jsg046>

Managing the driver distraction risk posed by technology: Changes to Australian road laws aimed at driver behaviour

Jeremy Wolter
Head of Legislative Reform
National Transport Commission
Level 3
600 Bourke Street, Melbourne, Victoria
AUSTRALIA
www.ntc.gov.au

Abstract:

Achieving national agreement and harmonisation on a technology neutral approach to road use policy in the Australian Federal system of government presented a unique set of challenges within a complex stakeholder environment.

Informed by the work of the National Transport Commission, Australian transport ministers have now approved a national policy- and subsequent amendments to the Australian Road Rules (ARR) – to address all sources of distraction while driving and provide a technology neutral approach to regulating interactions with technology. The policy is designed to manage the distraction risks posed by technology while encouraging innovation and ensuring technology that has the potential to improve safety is not constrained. The agreed policy encompasses a hybrid approach, using both prescriptive and performance-based rules, to clarify safe and unsafe interactions with technology. The policy intent is to reduce the road crash risk associated with driver distraction and provide better outcomes for road users regardless of the technology used.

The policy addresses the problem in the model Australian Road Rules (ARRs), through which uniform rules are prescribed for all road users across Australia, about the lack of clearly identified distracting activities that affect driving performance.

Amendments to the ARRs deliver on the national policy. Broadly the amendments include-

- **a broad prohibition to use technology (while the vehicle is moving or stationary but not parked), with lower risk interactions permitted by exception with inbuilt and mounted devices and motorcycle helmets**
- **prohibition of all physical interactions and restriction of visual interactions with wearable devices**
- **prohibition of all visual and physical interactions with non-mounted portable devices**

The national amendments are currently being adopted within state and territory transport legislation across Australia.

1. Introduction

Since 1999 the Australian Road Rules (ARRs) relating to driver distraction have been the basis for regulatory instruments used in Australia to deter unsafe driver engagement with secondary tasks while driving.

The ARRs regulate a broad range of sources of distraction that impact a driver's ability to maintain proper control and at the same time focus on specific types of technology that cause driver distraction, rather than on distracted driving behaviours and interactions that are known to be most risky from a safety perspective. The specific rules only preclude or limit the use of specific technology devices – mobile phones, visual display units and television receivers – while permitting their use as driver aids.

The ARRs –

- have not kept pace with the arrival of the smartphone and modern technology devices (including those built into the vehicle)
- inconsistently treat the sources of distraction and safety risks associated with certain behaviours
- can be confusing for road users and police regarding what technology devices are legal and illegal to use when driving.

The new Australian approach expands on the status quo, improves the ability to address all sources of distraction and is not limited to interactions with technology. It requires that drivers must ensure safe execution of non-driving-related tasks.

Technology neutral rules in road use regulatory settings for driver distraction in Australia provide an opportunity to encourage innovation and ensure that technology with the potential to improve road safety is not prohibited. Emerging transport technologies can provide opportunities to improve transport productivity and reduce deaths and injuries.

The approved Australian national policy framework for driver distraction focuses on driver behaviour and targets high-risk interactions with technology that are proven to significantly increase crash risk while driving.

It promotes the safe use of technology to operate a vehicle, conduct the professional driving task and navigate with a preference for lower risk audio and voice communication functionalities.

The new ARRs approved by all Australian governments prescribe rules for using electronic devices when driving. The rules cover how a person can interact with devices, what a person can use a device for and what the device may display on a screen. There are different rules depending on the device being used including inbuilt and mounted devices, motorbike helmet devices, wearable devices and portable devices.

Achieving national agreement and harmonisation in the area of road use policy in the Australian Federal system of government, however, presented a unique set of challenges. Each state and territory of Australia has near exclusive responsibility for road transport policy regulation.

Constitutionally, the Australian Commonwealth government has no direct powers of administration in this area unless incidental to another power conferred by the Australian constitution.

Another significant challenge was the need to ensure enforceability of new rules and clear community education about what is and is not allowed with respect to interaction with technology.

The agreed national policy framework and associated legislation represents three years of complex reform consultation and negotiation to produce a set of model rules suitable for implementation across Australia.

2. Australian Policy Approach/Method

The Australian policy approach was underpinned by five broad based principles as follows:

- An overarching requirement for a driver to have proper control of a vehicle to encourage safe use of technology regardless of whether an interaction is prohibited or not.
- Prescriptive rules for drivers are easily understood by the public and law enforcement agencies
- Prescriptive rules for drivers apply to all technology devices capable of wireless communication, electronic data retrieval, and/or displaying electronic data by display or projection

- Prescriptive rules for drivers apply to device interactions and functionalities known to result in an increased crash risk
- Voice-based interactions are permitted. There are no restrictions on voice-based interactions so long as the display is not visible to the driver in the normal driving position.

Based on these principles the national policy offers a suite of regulatory and non-regulatory approaches, including:

- A performance-based road rule – a tool to address both the observable driver and vehicle behaviours that cause and/or indicate the driver’s lack of control of a vehicle whether or not the source of the lack of control is based on a driver’s interaction with technology.
- Prescriptive road rules – introduction of four new device categories with a short, specific set of permitted and prohibited driver activities with technology addressing high-risk interactions to clarify what the public can and cannot do safely while driving.
- Non-regulatory tools – There are non-regulatory initiatives across the Australian transport system that will support the effectiveness of the changes to the ARRs. These include a safe driving guideline, public education campaigns and nationally-consistent messaging to ensure a shared and consistent understanding about the responsibilities of drivers in relation to driver distraction as well as the intent of the new road rules. This would capture, for example, the obligation on the driver to keep a proper lookout by paying attention to the surrounding road conditions and being able to intervene if required.

The four device categories targeted by the policy include inbuilt/mounted technology, wearables, portables and motorcycle/bicycle helmets.

The prescriptive element in the policy approach aims to encourage the take-up of new technologies (such as enhanced voice-user interfaces) consistent with Australian transport ministers’ priority to remove barriers to innovation and embrace new and emerging technologies.

3. The Governance and Stakeholder Environment

Since 2003, through an Intergovernmental Agreement (IGA), Australian States and Territories and the Commonwealth government have committed to improving transport productivity, efficiency, safety and environmental performance and regulatory efficiency in a uniform or nationally consistent manner.

The National Transport Commission (NTC) is responsible for development, maintenance and negotiation of transport related laws as tasked by transport ministers and departments of state. The NTC acts as a body independent of any one

Australian jurisdiction to practically deliver legislation outcomes which removes perceptions of conflict between Australian State and Commonwealth priorities.

Australian governments expect the NTC alongside Australian jurisdictions to lead core law reform. The commitment of Australian governments includes developing nationally consistent regulatory reform arrangements for road transport through the work of the NTC. Where appropriate, these law reforms may be expressed as model legislation so that consistency is promoted and maintained. The legislation for the driver distraction rules is prescribed in this way, and while they have no legal force in and of themselves, it is expected that the parties to the IGA will implement the agreed model legislation into local transport laws so that lawful enforceability and compliance is achieved.

Approval of model legislation requires unanimous agreement of all parties to the IGA. Achieving agreement with respect to national model uniform legislation presents many challenges because of the need to navigate the tension between the desire for a national response to emerging problems and the need to respect the constitutional separation of legislative powers between the Commonwealth and State and Territory jurisdictions.

4. Enforcement Challenges

Before the approval of the national policy and model legislation, the ambiguity of the rules for driver distraction made it difficult for enforcement agencies to identify behaviours that could result in distraction thereby reducing the driver distraction rules' safety benefits.

Australia's new vehicle market is small and therefore vehicle manufacturer's decisions about in-vehicle technologies has a direct effect on the potentially distracting features available to Australian motorists to use while vehicles are in motion or stationary but not parked. This means that enforcement of the road rules is one of the main regulatory tools to minimise driver distraction.

Achieving clarity in the new driver distraction rules was key to ensuring support at the national level. It was essential to ensure to the greatest extent possible the ability for enforcement to determine the applicable rule to the observed driver behaviour and therefore improve enforcement's likelihood to withstand scrutiny if questioned in a court of law.

The use of prescriptive rules seeks to facilitate enforcement by reducing the level of judgement enforcement officers exercise when applying the new rules. In combination with the use of prescriptive rules, the new rules include a performance-based element to target both the causes and consequences of driver engagement in distracting activities generally, regardless of whether they are technology based and not explicitly prohibited by law. The performance-based element is designed to target the effects of distracting activities, as well as the sources of distraction prior to a crash. The approach aims to mitigate the consequences of a wide

range of sources of distraction regardless of whether they are technology-based. A driver's engagement in non-technology-based activities, such as eating or attending to personal hygiene, may cause a driver to drive in a manner determined as failing to have proper control of a vehicle.

Finally, the ability for the new rules to be sufficiently enforceable through photo evidence to align with Australian States and territories automated camera enforcement of illegal mobile phone use was a key priority.

5. Australian Driver Distraction Model Laws

The Model laws approved by all Australian governments prescribe rules for using electronic devices when driving. The rules cover how a person can interact with devices, what a person can use a device for and what the device may display on a screen. There are different rules depending on the device being used (e.g mobile phone, smart watch or a vehicle infotainment system). For all electronic devices, however, the following interactions are prohibited:

- Typing of text or numbers into the device
- Scrolling through any content that is shown on a device's display
- Playing of movies, television shows, video games, animations, or other video content on a device that the driver can view
- Reading of text messages, group chats, emails or viewing of websites on the device
- Looking at social media or making video calls on the device.

The following summarises the rules for the different devices captured by the approved national policy. The rules provide a limited list of permitted interactions with technology, based on those interactions found by research to carry a lower risk of crash, including driver assistance functions such as navigation.

Visual and manual interactions found to carry a higher risk are consistently addressed through a broad prohibition to use technology (while the vehicle is moving or stationary but not parked), with lower risk interactions permitted by exception. It is permitted, however, to touch a device to stop an activity that is prohibited.

This approach is applied as consistently as practicable across four device categories which is a departure from the status quo and indicates what drivers can and cannot do with specific devices.

Inbuilt and mounted devices

Drivers must not touch an inbuilt or mounted device to type text or scroll through what is shown on the screen.

Drivers may interact with an inbuilt or mounted device in certain circumstances, some examples include, operating driving and vehicle systems, making a phone call, using navigation, playing music or other audio.

While a driver is permitted to touch an inbuilt or mounted device, the rules have been developed to only allow short interactions (for example, a single touch to select an option). If a driver needs to use the device for more than a moment, they will need to wait until they can safely pull over and legally park the vehicle.

It is not intended to penalise a driver for prohibited functions enabled by vehicle manufacturers that the driver of a vehicle has no control over.

Motorbike helmet devices

Motorbike riders must not touch a helmet device to type text or scroll through what is shown on a screen.

Riders may interact with the helmet device in certain circumstances, some examples include, , making a phone call, using navigation, playing music or other audio.

Wearable devices

Users of wearable devices must not touch the device to use its apps or functions. However, voice commands can be used to operate the device.

The wearable device can only show the user content about the safety and operation of the vehicle, making a phone call, or playing music.

Navigation and map functions on the device must not be visible to the driver in the normal driving position.

Portable devices

Drivers must not touch or hold a portable device regardless of whether it is on or off.

Drivers must not be able to see anything on the portable device screen from the normal driving position apart from automatic notifications and basic information such as time, date and battery power.

Drivers can use the portable device with hands-free controls provided the driver cannot see the screen from the normal driving position.

6. Conclusion

The agreed Australian position on regulation of driver distraction through law includes a combination of performance-based and prescriptive rules.

This combination provides:

- A clear indication of permitted and prohibited interactions with technology based on high-risk interactions and behaviours identified by research.
- A performance-based component that addresses any sources of distraction that could impair a driver's proper control of a vehicle and clear view of the road and traffic.

It is expected that this approach will provide the highest road safety-benefits in terms of reducing the number of fatalities, injuries and economic costs from accidents.

In addition, there are non-regulatory initiatives across the Australian transport system that will support achieving the overall policy objective and enhance the effectiveness of the new model driver distraction rules.

7. References

- ABS 2019, *9208.0 – Survey of Motor Vehicle Use, Australia, 12 months ended 30 June 2018*, accessed at: <https://www.abs.gov.au/ausstats/abs@.nsf/mf/9208.0>
- AT&T 2015, *Smartphone Use While Driving Grows Beyond Texting to Social Media, Web Surfing, Selfies, Video Chatting*, 19 May 2015, accessed at: http://about.att.com/story/smartphone_use_while_driving_grows_beyond_texting.html
- Attorney-General's Department 2019, *Reducing the complexity of legislation*, accessed at: <https://www.ag.gov.au/LegalSystem/ReducingTheComplexityOfLegislation/Pages/default.aspx>
- Australasian New Car Assessment Program (ANCAP) 2019, *ANCAP Submission: Addressing NTC Issues Paper on Developing Technology-Neutral Road Rules for Driver Distraction*, February 2019, accessed at: http://www.ntc.gov.au/system/files/webform/submission_dd_ip/187/Developing-technologyneutral-road-rules-for-driver-distraction-submission-ancap-safety.pdf
- Australian Light Vehicle Standards Rules, as at 22 November 2019.
- Australian Mobile Telecommunications Association 2019, *AMTA Submission to National Transport Commission Project: Developing technology-neutral road rules for driver distraction*, February 2019, accessed at: http://www.ntc.gov.au/system/files/webform/submission_dd_ip/181/Developing-technologyneutral-road-rules-for-driver-distraction-submission-amta-1.pdf
- Australian Road Rules, as at 19 March 2018.
- Ba, Y, Zhang, W, Peng, Q, Salvendy, G & Crundall, D 2015, Risk-taking on the road and in the mind: behavioural and neural patterns of decision making between risky and

safe drivers, *Ergonomics*, July 2015, accessed at: <http://dx.doi.org/10.1080/00140139.2015.1056236>

Beanland, V, Fitzharris, M, Young, K & Lenné, MG 2013, Driver inattention and driver distraction in serious casualty crashes: data from the Australian National Crash In-Depth Study, *Accident Analysis & Prevention*, vol. 54, pp. 99–107, accessed at <https://www.sciencedirect.com/science/article/abs/pii/S000145751300047X>

Birrell, SA & Young, MS 2011, The impact of smart driving aids on driving performance and driver distraction, *Transportation Research, Part F*, no. 14, pp. 484–493, accessed at <https://www.sciencedirect.com/science/article/abs/pii/S1369847811000751>

CCS Insight 2019, *Optimistic outlook for wearables: 260 Million Unit Sales in 2023*, accessed at: <https://www.ccsinsight.com/press/company-news/optimistic-outlook-for-wearables/>

Centre for Accident Research & Road Safety – Queensland, 2017, *Scoping Study of Mobile Phone Use While Driving – final report*, accessed at: <http://eprints.qut.edu.au/121618/1/121618.pdf>

Cervantes, E 2019, Best smart motorcycle helmets to smarten up your ride, *Android Authority*, 6 August 2019, accessed at: <https://www.androidauthority.com/best-smartmotorcycle-helmets-1014050/>

Chevalier, A, Cunningham, M & Roberts, P 2019, *ARRB response to the NTC's Driver Distraction Issues Paper*, Australian Road Research Board (ARRB), accessed at: [https://www.ntc.gov.au/Media/Reports/\(D2686DC6-21DA-4139-888C-353B6BB9B96\).pdf](https://www.ntc.gov.au/Media/Reports/(D2686DC6-21DA-4139-888C-353B6BB9B96).pdf)

Christie, R & Harrison, W 2003, *Driver training and education programs of the future*, Report No. 03/03, RACV Ltd, Melbourne

Commissioner for Better Regulation 2016, *Victorian Guide to Regulation*, accessed at: <https://www.vic.gov.au/sites/default/files/2019-10/Victorian-Guide-to-Regulation.pdf>

Commonwealth of Australia, 2018a, *National Road Safety Strategy 2011–2020*, Australian Government, accessed at: <http://roadsafety.gov.au/nrss/>

Commonwealth of Australia, 2018b, *The Safe System approach*, Australian Government, accessed at: <http://roadsafety.gov.au/nrss/safe-system.aspx>

COAG 2007, *A Guide for Ministerial Councils and National Standard Setting Bodies*, Best Practice Regulation, October 2007, accessed at: https://www.pmc.gov.au/sites/default/files/publications/C OAG_best_practice_guide_2007.pdf

Cunningham, ML & Regan, MA 2015, *Taxonomy of Driver Distraction*, Australian Road Research Board, Sydney (report for VicRoads).

Deloitte 2018, *Behaviour unlimited: Mobile Consumer Survey 2018 – The Australian cut*, accessed at: <https://www2.deloitte.com/au/mobile-consumer-survey>

Department of Infrastructure, Regional Development and Cities 2018, *Community attitudes to road safety – 2017 survey report*, accessed at: https://infrastructure.gov.au/roads/safety/publications/2018/pdf/community_att_17.pdf

Department of Transport and Main Roads (Queensland) 2019, *Submission in response to the National Transport Commission's consultation Regulation Impact Statement on developing technology-neutral road rules for driver distraction*, accessed at: http://www.ntc.gov.au/system/files/webform/submission_distraction_ris/663/Developingtechnology-neutral-road-rules-for-driver-distraction-RIS-submission-TMR.pdf

Dingus, TA, Guo, F, Lee, S, Antin, JF, Perez, M, Buchanan-King, M & Hankey, J 2016, Driver crash risk factors and prevalence evaluation using naturalistic driving data, *Proceedings of the National Academy of Sciences*, vol. 113, no. 10, pp. 2636–2641.

DriveRisk Australasia 2019, *Submission: Developing technology-neutral road rules for driver distraction*, February 2019, accessed at: http://www.ntc.gov.au/system/files/webform/submission_dd_ip/180/Developing-technologyneutral-road-rules-for-driver-distraction-submission-owen-neochi.pdf

Ehsani, J, Bingham, C, Ionides, E & Childers, D 2014, The impact of Michigan's text messaging restriction on motor vehicle crashes, *Journal of Adolescent Health*, vol. 54, pp. S68–S74, accessed at: https://www.researchgate.net/profile/C_Raymond_Bingham/publication/261837898_The_Impact_of_Michigan's_Text_Messaging_Restriction_on_Motor_Vehicle_Crashes/links/5653044108ae619b18ec8c/The-Impact-of-Michigans-Text-Messaging-Restriction-on-Motor-Vehicle-Crashes.pdf

Elder, RW, Shults, RA, Sleet, DA, Nichols, JL, Thompson, RS & Rajab, W 2004, Effectiveness of mass media campaigns for reducing drinking and driving and alcohol involved crashes: a systematic review, *American Journal of Preventive Medicine*, vol. 27, no. 1, July 2004, pp. 57–65, accessed at: <https://www.sciencedirect.com/science/article/abs/pii/S0749379704000467>

Elliot, B 1992, *Report on achieving high levels of compliance with road safety laws – a review of road user behaviour modification*, Report No. 6, Brisbane: Travelsafe Committee, Legislative Assembly of Queensland.

European Road Safety Observatory 2015, *Driver distraction*, accessed

at:https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/ersosynthesis2015-detail-driverdistraction25_en.pdf

FCAI 2019a, *Written Submission in Response to Regulation Impact Statement – Developing Technology – Neutral Road Rules for Driver Distraction*, accessed at:http://www.ntc.gov.au/system/files/webform/submission_distraction_ris/443/Developingtechnology-neutral-road-rules-for-driver-distraction-RIS-submission-FCAI.pdf

FCAI 2019b, *Australia's new vehicle market*, accessed at:<https://www.fcai.com.au/sales>

Ferdinand, A, Menachemi, N, Sen, B, Blackburn, J, Morrisey, M & Nelson, L 2014, Impact of texting laws on motor vehicular fatalities in the United States, *American Journal of Public Health*, vol. 104, no. 8, pp. 1370–1377, accessed at:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4103220/pdf/AJPH.2014.301894.pdf>

Ferdinand, A, Menachemi, N, Blackburn, J, Sen, B, Nelson, L & Morrisey, M 2015, The impact of texting bans on motor vehicle crash-related hospitalizations, *American Journal of Public Health*, vol. 105, no. 5, pp. 859–865, accessed at:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4386499/pdf/AJPH.2014.302537.pdf>

Financial Times 2018, *Australia's Uber drivers earn less than minimum wage*, accessed at:<https://www.ft.com/content/f9ce2cd8-d5c2-11e8-a854-33d6f82e62f8>

Fitch, GM, Soccolich, SA, Guo, F, McClafferty, J, Fang, Y, Olson, RL, Perez, MA, Hanowski, RJ, Hankey, JM & Dingus, TA 2013, *The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk*, National Highway Traffic Safety Administration, Washington, DC: Rep. No. DOT HS 811 757

Fitzharris, M, Young, K & Bowman, D 2012, *Potential benefits of 'driver distraction' regulatory reform*, Monash University Accident Research Centre, Clayton Funkhouser, D & Chrysler, ST 2007, *Assessing Driver Distraction Due to In-Vehicle Video Systems through Field Testing at the Pecos Research and Testing Center*, Report No. SWUTC/07/473700-00082-1, Southwest Region University Transportation Center, College Station, TX, 2007, accessed at:<https://static.tti.tamu.edu/swutc.tamu.edu/publications/technicalreports/473700-00082-1.pdf>

General Motors 2019, *Marketplace*, accessed at:<https://gmmarketplace.gm.com/>

Giang, WCW, Shanti, I, Chen, H-YW, Zhou, A & Donmez, B 2015, *Smartwatches vs. smartphones: a preliminary report of driver behavior and perceived risk while responding to notifications*, Proceedings of the 7th International Conference on Automotive User Interfaces

and Interactive Vehicular Applications, Nottingham, United Kingdom, Association for Computing Machinery, New York, pp. 154–161, accessed at:https://www.researchgate.net/profile/Wayne_Chi_Wei_Giang/publication/280024357_Smartwatches_vs_Smartphones_A_Preliminary_Report_of_Driver_Behavior_and_Perceived_Risk_while_Responding_to_Notifications/links/5626c0ea08ae4d9e5c4d4699/Smartwatches-vsSmartphones-A-Preliminary-Report-of-Driver-Behavior-and-Perceived-Risk-while-Responding-to-Notifications.pdf

Goodsell, R, Cunningham, M & Chevalier, A 2019, *Driver Distraction: A Review of Scientific Literature*, Project No. 013817, prepared for the National Transport Commission, accessed at:<https://www.ntc.gov.au/sites/default/files/assets/files/Driver%20distraction%20-%20A%20review%20of%20scientific%20literature.pdf>

Hartley, J 2007, Driver distraction: a law enforcement perspective. In: Faulks, Regan, Stevenson, Brown, Porter & Irwin (eds.). *Distracted driving*, Australasian College of Road Safety, Sydney, pp. 329–344, accessed at:<http://acrs.org.au/files/papers/13%20Hartley%20A%20law%20enforcement%20perspective.pdf>

Hickman, JS, Hanowski, RJ & Bocanegra, J 2010, *Distraction in commercial trucks 476 and buses: Assessing prevalence and risk in conjunction with crashes and near-crashes*, Federal Motor Carrier Safety Administration, Washington, DC: Rep. No. FMCSA-RRR-010-049

Highway Loss Data Institute, 2010, Texting laws and collision claim frequencies, *Highway Loss Data Institute Bulletin*, vol. 27, no. 11, pp. 1–10, accessed at:https://www.iihs.org/media/fc495300-6f8c-419d-84d7-c3b94d178e5a/enPLrA/HLDI%20Research/Bulletins/hldi_bulletin_27.11.pdf

IBISWorld 2019a, *Courier Pick-up and Delivery Services*, Industry Report I5102

IBISWorld 2019b, *Ridesharing Services in Australia*, Industry Report OD5540

Insurance Australia Group 2019, *Developing technology-neutral road rules for driver distraction*, February 2019, accessed at:http://www.ntc.gov.au/system/files/webform/submission_dd_ip/169/Developing-technology-neutral-road-rules-for-driver-distraction-submission-iag.pdf

Isler, RB, Starkey, NJ & Sheppard, P 2011, Effects of higher-order driving skill training on young, inexperienced drivers' on-road driving performance, *Accident Analysis and Prevention*, no. 43, pp. 1818–1827

Johnson, SB & Jones, VC 2011, Adolescent development and risk of injury: using developmental science to improve interventions, *Injury Prevention*, no. 17, pp. 50–54

- Kemp, A & Gu, H 2017, *Use of ride sharing and taxi services in New South Wales – A report for the Independent Pricing and Regulatory Tribunal*, Houston Kemp, November 2017, accessed at: <https://www.ipart.nsw.gov.au/files/sharedassets/website/shared-files/pricingreviews-transport-services-admin-taxi-fares-and-licences-from-july-2018/publications-taxifares-and-licences-from-july-2018/consultant-report-houstonkemp-use-of-ride-sharing-and-taxi-services-in-nsw-december-2017.pdf>
- Klauer, SG, Dingus, TA, Neale, VL, Sudweeks, JD, & Ramsey, DJ 2006, *The impact of driver inattention on near-crash/crash risk: an analysis using the 100-Car Naturalistic Driving Study data*, National Highway Traffic Safety Administration, DOT HS 810 594, April 2006, accessed at: <https://vtechworks.lib.vt.edu/bitstream/handle/10919/55090/DriverInattention.pdf>
- Kwon, O, Yoon, Y & Jang, K 2014, Evaluating the effectiveness of the law banning handheld cellphone use while driving, *Safety Science*, vol. 70, pp. 50-57, accessed at: <https://www.sciencedirect.com/science/article/pii/S092575351400099X>
- Lassa, T 2019, *U.S. Auto Sales Totaled 17.25-Million in 2017*, accessed at: <https://www.automobilemag.com/news/u-s-auto-sales-totaled-17-25-million-calendar-2017/>
- Legislative Standards Act 1992* (Qld), section 4(3)(k).
- McKnight, AJ & Bahouth, GT 2008, Analysis of large truck rollover crashes, *Association for the Advancement of Automotive Medicine*, vol. 52, pp. 281–288, accessed at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3256782/>
- Motorcycle Council of New South Wales 2019, *National Transport Commission – Developing technology-neutral road rules for driver distraction – Consultation Regulation Impact Statement*, August 2019, accessed at: http://www.ntc.gov.au/system/files/webform/submission_distraction_ris/440/Developingtechnology-neutral-road-rules-for-driver-distraction-RIS-submission-Motorcycle-Council-of-NSW.pdf
- NHTSA 2019, *Traffic safety facts research note: Distracted Driving in Fatal Crashes, 2017*. Research note, DOT HS 812 700, accessed at: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812700>
- NHTSA 2018, *Traffic safety facts research note: Distracted Driving 2016*. Research note, DOT HS 812 517, accessed at: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812517>
- NatRoad 2019a, *Submission to the National Transport Commission: Developing technology-neutral road rules for driver distraction – Issues Paper*, February 2019, accessed at: http://www.ntc.gov.au/system/files/webform/submission_distraction_ris/189/Developingtechnologyneutral-road-rules-for-driver-distraction-submission-natroad.pdf
- NatRoad 2019b, *Submission to the National Transport Commission: Developing technology-neutral road rules for driver distraction – Consultation Regulation Impact Statement*, August 2019, accessed at: http://www.ntc.gov.au/system/files/webform/submission_distraction_ris/439/Developingtechnology-neutral-road-rules-for-driver-distraction-RIS-submission-Natroad.pdf
- National Transport Commission 2017a, *Changing driving laws to support automated vehicles – discussion paper*, October 2017, accessed at: [https://www.ntc.gov.au/Media/Reports/\(E5695ACE-993C-618F-46E1-A876391B8CD9\).pdf](https://www.ntc.gov.au/Media/Reports/(E5695ACE-993C-618F-46E1-A876391B8CD9).pdf)
- National Transport Commission 2017b, *Work Program 2017–21*, accessed at: [https://www.ntc.gov.au/Media/Reports/\(5221BAAC-9DAB-B48F-54F2-C96CFD60D3CE\).pdf](https://www.ntc.gov.au/Media/Reports/(5221BAAC-9DAB-B48F-54F2-C96CFD60D3CE).pdf)
- OICA 2015, *Recommended OICA Worldwide Distraction Guideline Policy Position*, March 2015, accessed at: <http://www.oica.net/wp-content/uploads/OICA-Position-Paper-Driver-Distraction-Final-2015-03-03.pdf>
- OBPR 2007, *Best Practice Regulation: A guide for Ministerial Councils and National Standard Setting Bodies*, accessed at: https://www.pmc.gov.au/sites/default/files/publications/C_OAG_best_practice_guide_2007.pdf
- Olson, RL, Hanowski, RJ, Hickman, JS & Bocanegra, J 2009, *Driver distraction in commercial vehicle operations*, Report No. FMCSA-RRR-09-042, US Department of Transportation, Washington, D.C., 2009, accessed at: <https://rosap.nhtsa.gov/viewdot/17715>
- Paine, M & Regan, MA 2018, *Road Map for a Human Machine Interface Distraction Safety Rating*, Australian Road Research Board, Sydney (report for VicRoads)
- Parnell, KJ, Stanton, NA & Plant, KL 2018, What technologies do people engage with while driving and why?, *Accident Analysis and Prevention*, no. 111, pp. 222–237
- Parnell, KJ, Stanton, NA & Plant, KL 2017, What's the law got to do with it? Legislation regarding in-vehicle technology use and its impact on driver distraction, *Accident Analysis and Prevention*, no. 100, pp. 1–14, accessed at: <https://www.sciencedirect.com/science/article/abs/pii/S001457516304535>
- Regan, M, Hallett, C & Gordon, CP 2011, Driver distraction and driver inattention: definition, relationship and taxonomy, *Accident Analysis and Prevention*, no. 43, pp. 1771–1781, accessed at: https://www.cambridge.org/core/Regan_et_al_2011_driver_distraction.pdf

- Regan, M & Prabhakaran, P 2019, *Submission to National Transport Commission in Response to Issued Paper Developing Technology-neutral Road Rules for Driver Distraction*, Research Centre for Integral Transport Innovation, University of New South Wales Sydney, accessed at: [https://www.ntc.gov.au/Media/Reports/\(A352B22F-1424-085D-76EC-24A6D33508AC\).pdf](https://www.ntc.gov.au/Media/Reports/(A352B22F-1424-085D-76EC-24A6D33508AC).pdf)
- Regan, M, Prabhakaran, P & Dixit, V 2019, *Submission to National Transport Commission in Response to Consultation Regulation Impact Statement Developing Technology-neutral Road Rules for Driver Distraction*, Research Centre for Integral Transport Innovation, University of New South Wales Sydney, accessed at: http://www.ntc.gov.au/system/files/webform/submission_distraction_ris/655/Developingtechnology-neutral-road-rules-for-driver-distraction-RIS-submission-transport-for-RCITI.pdf
- Regan, M & Young, K 2007, Driver distraction: a review of the literature, In: I.J. Faulks, M. Regan, M. Stevenson, J. Brown, A. Porter & J.D. Irwin (Eds.) *Distorted driving*, Sydney, NSW: Australasian College of Road Safety, pp. 379-405, accessed at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.469.4178&rep=rep1&type=pdf>
- Revised Code of Washington 2019*, Title 46, Chapter 46.61, Section 46.61.672, Washington State, United States of America, accessed at: <https://app.leg.wa.gov/RCW/default.aspx?cite=46.61.672>
- Road Safety Commission 2018, *2017 Preliminary summary of fatalities on Western Australian roads*, Government of Western Australia, accessed at: <https://www.rsc.wa.gov.au/RSC/media/Documents/Road%20Data/Statistics/Annual%20crash%20statistics/annual-prelim-crash-statistics-2017.pdf>
- Roper, J 2017, *A Study of the Effectiveness of the Hands-Free Ordinance in San Antonio, Texas*, accessed at: <https://digital.library.txstate.edu/bitstream/handle/10877/6626/RoperJacob.pdf?sequence=1&isAllowed=y>
- Royal Automobile Association of South Australia 2019, *Developing Technology Neutral Road Rules for Driver Distraction: RAA submission to the National Transport Commission*, February 2019, accessed at: http://www.ntc.gov.au/system/files/webform/submission_dd_ip/177/Developing-technologyneutral-road-rules-for-driver-distraction-submission-raa.pdf
- Royal Automobile Club of Victoria 2019, *Submission to Issues Paper for Developing Technology-Neutral Road Rules for Driver Distraction*, February 2019, accessed at: http://www.ntc.gov.au/system/files/webform/submission_dd_ip/173/Developing-technologyneutral-road-rules-for-driver-distraction-submission-racv.pdf
- Rudisill, TM, Chu, H, Zhu, M 2018, *Cell phone use while driving laws and motor vehicle driver fatalities: differences in population subgroups and location*, *Annals of Epidemiology* vol. 28, pp. 730-735.
- Strayer, DL 2015, Is the technology in your car driving you to distraction?, *Policy Insights From the Behavioral and Brain Sciences*, vol. 2, no. 1, pp. 157–165, accessed at: https://www.researchgate.net/profile/David_Strayer2/publication/281479990_Is_the_Technology_in_Your_Car_Driving_You_to_Distraction/links/561bbe7f08aea80367242d2a/Is-the-Technology-in-Your-Car-Driving-You-to-Distraction.pdf
- Strayer, DL, Cooper, JM, Turrill, J, Coleman, JR & Hopman, J 2016, Talking to your car can drive you to distraction, *Cognitive Research: Principles and Implications*, vol. 1, November 2016, accessed at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5256443/>
- Strayer, DL, Cooper, JM, Goethe, RM, McCarty, MM, Getty, D & Biondi, F 2017, *Visual and cognitive demands of using in-vehicle infotainment systems*, Washington, DC: AAA Foundation for Traffic Safety, accessed at: <https://aaafoundation.org/wpcontent/uploads/2017/11/VisualandCognitive.pdf>
- Strayer, DL, Cooper, JM, McCarty MM, Getty, DJ, Wheatley, CL, Motzkus, CJ, Mackenzie, KL, Loveless, SM, Esplin, J, Goethe, RM & Biondi, F 2018, *Visual and cognitive demands of using Apple's CarPlay, Google's Android Auto and five different OEM infotainment systems*, AAA Foundation for Traffic Safety, accessed at: <https://aaafoundation.org/wpcontent/uploads/2018/06/AAA-Phase-6-CarPlay-Android-Auto-FINAL.pdf>
- Strayer, DL, Cooper, JM, Goethe, RM, McCarty, MM, Getty, DJ & Biondi, F 2019, Assessing the visual and cognitive demands of in-vehicle information systems, *Cognitive Research: Principles and Implications*, vol. 4, no. 18, accessed at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6588669/>
- Tesla 2019, *Introducing Software Version 10.0*, 26 September 2019, accessed at: https://www.tesla.com/en_AU/blog/introducing-software-version-10-0
- Transport Accident Commission 2016, *Towards Zero 2016–2020, Victoria's Road Safety Strategy and Action Plan*, accessed at: https://www.towardszero.vic.gov.au/_data/assets/pdf_file/0010/183556/STU_0206_RS_STRATEGY_2016_web.pdf
- Transport Accident Commission 2018, *Young driver statistics*, accessed at: <http://www.tac.vic.gov.au/road-safety/statistics/summaries/young-driver-statistics>
- Transport and Infrastructure Council 2019, *Transport and Infrastructure Council Strategic Work Programme*, accessed at: https://www.transportinfrastructurecouncil.gov.au/publications/files/Council_Strategic_Work_Programme.pdf

- Transport for NSW 2019, *Transport for NSW Submission*, accessed at: http://www.ntc.gov.au/system/files/webform/submission_distraction_ris/434/Developing-technology-neutral-road-rules-for-driver-distraction-RIS-submission-transport-fornsw.pdf Developing technology-neutral road rules for driver distraction October 2020
- Transport for NSW 2020, Mobile phone detection cameras, accessed at: <https://roadsafety.transport.nsw.gov.au/stayingsafe/mobilephones/technology.html>
- Transport Workers Union 2018, *Rideshare Driver Survey*, accessed at: <https://www.twu.com.au/Home/Campaigns/Rideshare-Divers/Rideshare-Survey-Infographic/>
- Tranter, P & Warn, J 2008, Relationships between interest in motor racing and driver attitudes and behaviour amongst mature drivers: An Australian case study, *Accident Analysis & Prevention*, vol. 40, no. 5, pp. 1683–1689
- Twisk, D 2007, *Trends in risk of young drivers and countermeasures in European Countries*, Paper presented to National Safety Council's International Symposium on Novice Teen Driving: GDL and Beyond, February 5–7, Tucson, AZ
- Uber Technologies Inc. 2019, *Common FAQs partners have before their first trip*, Pre-first trip: Frequently asked questions, accessed at: <https://www.uber.com/drive/new-orleans/resources/prefirst-trip-faqs/>
- Ulleberg, P & Rundmo, T 2003, Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers, *Safety Science*, vol. 41, no. 5, pp. 427–443
- Verschuur, WL & Hurts, K 2008, Modelling safe and unsafe driving behaviour, *Accident Analysis & Prevention*, vol. 40, no. 2, pp. 644–656
- The Washington Post 2017, *A man using FaceTime killed a 5-year-old girl in a highway crash. Was Apple to blame?*, 2 January 2017, accessed at: https://www.washingtonpost.com/news/the-intersect/wp/2017/01/02/a-man-using-facetime-killed-a-5-year-old-girl-in-a-highway-crash-was-apple-to-blame/?noredirect=on&utm_term=.f0398399739b
- Washington State Legislature 2017, *Distracted Driving – Electronic Devices – Dangerously Distracted Driving*, Substitute Senate Bill 5289, accessed at: <http://lawfilesexternal.wa.gov/biennium/2017-18/Pdf/Bills/Session%20Laws/Senate/5289-S.SL.pdf?cite=2017%20c%20334%20>
- Watson, B 1997, *When common sense just won't do: Misconceptions about changing the behaviour of road users*, In: Bullen & Troutbeck (eds), *The Second International Conference on Accident Investigation, Reconstruction, Interpretation & the Law: Proceedings, 20–23 October 1997* (pp. 347–359), Brisbane
- Watson, JM & Strayer, DL 2010, Supertaskers: Profiles in extraordinary multitasking ability, *Psychonomic Bulletin & Review*, vol. 17, no. 4, pp. 479–485
- Williamson, A, Hatfield, J & Friswell, R 2019, *Submission to the National Transport Commission regarding Developing technology-neutral road rules for driver distraction: Consultation regulation impact statement*, Transport and Road Safety (TARS) Research Centre, University of New South Wales, accessed at: https://www.ntc.gov.au/submission_data/454
- Yagil, D 2005, *Drivers and traffic laws: A review of psychological theories and empirical research*, Traffic and Transport Psychology, Elsevier, Oxford, accessed at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.528.6758&rep=rep1&type=pdf>
- Young, K, Horberry, T & Charlton, J 2019, *Submission to the National Transport Commission on the Issues Paper: Developing Technology-Neutral Road Rules for Driver Distraction*, Monash University Accident Research Centre (MUARC), accessed at: <https://www.ntc.gov.au/media/1938/ntc-issues-paper-developing-technology-neutral-road-rules-for-driver-distraction-kristie-young-monash-university-accident-research-centre-feb-2019.pdf>
- Young, K & Lenné, M, 2012, *Driver distraction regulatory reform in Australia*, Monash University Accident Research Centre, Clayton.
- Young, K, Osborne, R, Koppel, S, Charlton, J, Grzebieta, R, Williamson, A, Haworth, N, Woolley, J & Senserrick, T 2018, What are Australian drivers doing behind the wheel? An overview of secondary task data from the Australian Naturalistic Driving Study, *Journal of the Australasian College of Road Safety*, accessed at: <http://acrs.org.au/wp-content/uploads/What-are-Australian-drivers-doing-behind-the-wheel-An-overview-of-secondary-task-data-from-the-Australian-Naturalistic-Driving-Study.pdf>
- Young, K & Regan, M 2003, *Driver distraction: A review of the literature and recommendations for countermeasure development*, Proceedings of the 2003 Road Safety Research, Policing and Education Conference, 24–26 September, accessed at: <https://acrs.org.au/files/arsrpe/RS030132.pdf>
- Young, K & Regan, M 2007, 'Driver distraction: A review of the literature'. In: Faulks, Regan, Stevenson, Brown, Porter & Irwin (eds), *Distracted driving*, Australasian College of Road Safety, Sydney, pp. 379-405, accessed at: <http://acrs.org.au/files/papers/15%20Young%20A%20review%20of%20the%20literature.pdf>

Non-driving related task engagement in highly automated vehicles: How to mitigate emerging motion sickness?

Claus Marberger^{1*}, Michael Schulz², Philipp Alt², Michael Teicht³, Arnd Engeln³

¹ Corporate Sector Research and Advance Engineering, Robert Bosch GmbH, Robert-Bosch-Campus 1, 71272 Renningen, Germany

² Cross-Domain Computing Solutions, Robert Bosch GmbH, Robert-Bosch-Allee 1, 74232 Abstatt, Germany

³ Institute for Mobility and Digital Innovation, Hochschule der Medien, Nobelstr. 10, 70569 Stuttgart, Germany

* claus.marberger@de.bosch.com

In this study, three evidence-based countermeasures to mitigate motion sickness in automated vehicles have been compared with respect to a control condition. The measures were based on visual anticipatory cues for vehicle motion or on the optimized alignment of the human body by seat adjustments. Test subjects ($N = 28$) experienced each condition on a 20 minutes' drive in a highly automated vehicle on a closed test track. The non-driving related task was to read a text on a handheld tablet while being exposed to automated fore-aft movements (representing stop and go traffic conditions). None of the implemented countermeasures could be shown to significantly mitigate motion sickness under the circumstances of the study. The paper finishes by discussing methodological issues and possible confounding factors.

1. Introduction

Driving automation at SAE level 3 (SAE International, 2021) and above is expected to enable non-driving related tasks (NDRT) for all passengers. At the same time a significant percentage of users will be confronted with issues of motion sickness – a phenomenon which is commonly explained by a mismatch of sensed and expected motion stimuli (Reason, 1978). In order to reduce the amount of mismatch, countermeasures of various kinds have been proposed and evaluated by researchers, e.g., elevated display positions in order to preserve as much environmental awareness as possible (e.g. Kuipers, 2018; Brietzke, 2021), peripheral visual cues indicating upcoming turn manoeuvres (Karjanto et al., 2018), anticipatory audio cues preceding fore-aft vehicle motion (Kuipers et al., 2020), compensation of horizontal accelerations by tilting/moving seats (Golding et al., 2013, Donohew & Griffin, 2009) or permanently reclined seating positions (Bohrmann & Bengler, 2019). Published studies typically report a positive (yet subtle) impact of the inspected countermeasures on the emergence of motion sickness. However, it is difficult to compare the effects of the countermeasures across publications since studies vary in type and level of applied motion dose, the recruited test sample or the way motion sickness is measured. According to the opinion of the authors there is also a lack of publications that aim to replicate the effects found in original studies. The aim of this study was to compare three implementations of evidence-based countermeasures according to the state-of-the art and to compare their effectivity in a controlled experimental setting.

2. Method

2.1 Research apparatus and motion profile

The closed-track study was based on a research vehicle allowing to fully automate longitudinal and lateral control. The motion profile was designed to represent a continuous stop-and-go traffic scenario including frequent fore-aft

acceleration events with peak values of $+2 \text{ m/s}^2$ and -3 m/s^2 , respectively. The maximum speed was limited to 60 km/h and the duration of the drive was set to 20 minutes (see Fig. 1 for the longitudinal acceleration profile). Lateral accelerations were intentionally kept to a minimum.

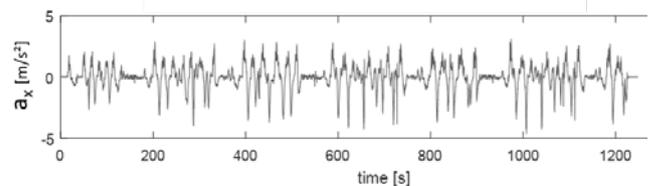


Fig 1: Longitudinal acceleration profile representing stop-and-go traffic conditions.

2.2 Test sample

A total of $N = 28$ participants (14 females, 14 males) took part in the study. Their age ranged from 23 to 47 years ($M = 35.61$; $SD = 7.62$). All participants were screened in advance for increased susceptibility to motion sickness based on self-assessment. The average score of the test sample on the MSSQ Short scale was 19.18 which represents the 75th percentile of the population (Golding, 2006). Informed consent was given by all participants in advance. Every subject could terminate the trials at any time without negative consequences.

2.3 Independent variables

The experimental study was based on a within-subject design with the following 4 conditions, presented on 4 separate days in counterbalanced order (see also Fig. 2). In all conditions, the participants were instructed to continuously read text on a handheld device.

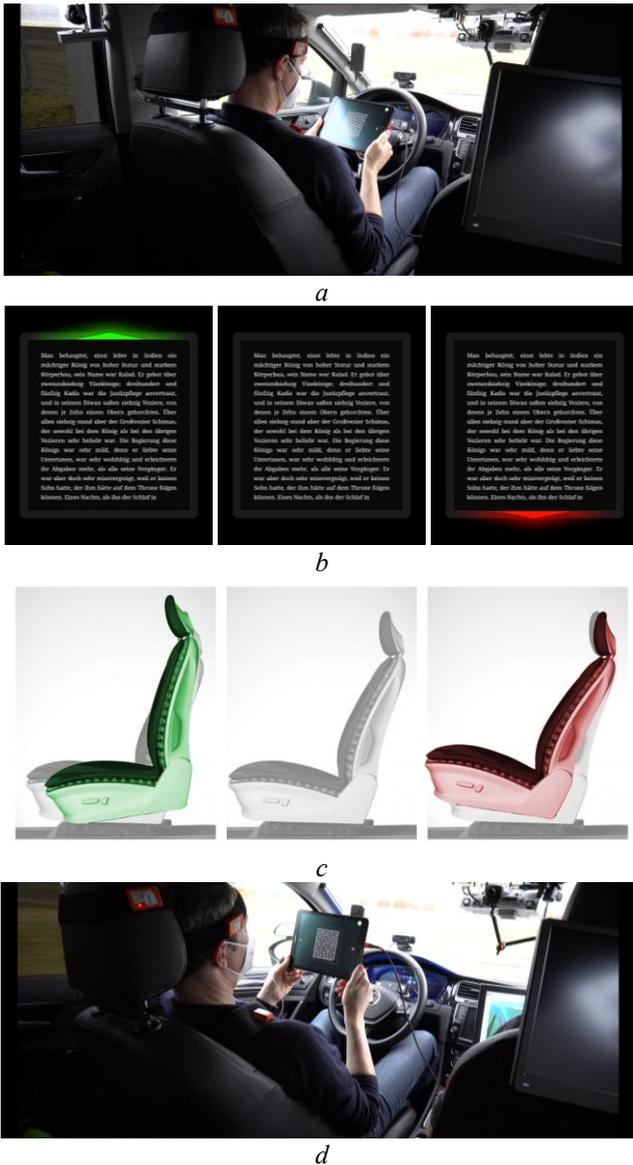


Fig. 2. Experimental conditions (see details in the text) (a) Control condition, (b) visual anticipatory cues, (c) dynamic seat adjustments, (d) permanently reclined seating

2.3.1 Control condition

The subject was seated on the driver’s seat behind the steering wheel in an upright sitting position (inclination angle of backrest: 25°; see Fig. 2a).

2.3.2 Visual anticipatory cues

The subject received dedicated visual cues below and above the text box on the handheld tablet (see Fig. 2b), preceding the actual onset of vehicle acceleration/deceleration by 1.3 seconds. The brightness of the visual cues was linked to the level of the upcoming acceleration (from transparent to full brightness). All other aspects were identical to the control condition.

2.3.3 Dynamic seat adjustment

In this condition the seat moved along a curved trajectory in longitudinal direction (cf. forward swing

movement for braking, backward swing movement for accelerating). This dynamic seat movement was implemented as a pre-programmed adjustment of three seat actuators (forward-backward, tilt and backrest angle). The resulting inclination angles for the backrest were 30° (neutral position), up to 20° for positive vehicle acceleration and up to 40° for negative vehicle acceleration (see Fig. 2c).

2.3.4 Permanently reclined seating position

Participants were instructed to read the text while the backrest of the seat was permanently set in a reclined position (inclination angle of backrest: 40°; see Fig. 2d). Subjects could use the headrest while reading on the handheld device.

2.4 Dependent variables

The main dependent (subjective) variable was based on differences between pre and post scores of the Motion Sickness Assessment Questionnaire (Delta MSAQ) according to Gianaros et al., 2001. The questionnaire was administered before and after all test drives.

2.5 Test procedure

Each of the four trials started with a pre-drive questionnaire and basic instructions, followed by a 20 minutes’ drive in the research vehicle. During the drive the participants were requested to continuously engage in the reading task. In order to monitor the development of motion sickness in real time, test subjects indicated their current level of motion sickness on a 10-point rating scale every 2 minutes. The safety driver sitting next to the participant was instructed to terminate the trial whenever the participant wishes or when the reported motion sickness level exceeded the value of 6. A post-drive questionnaire with interview followed in the office next to the test track location.

3. Results

3.1 General effects on motion sickness mitigation

In order to compare the level of emerged motion sickness between the experimental conditions, the distribution of Delta MSAQ scores is analysed (see Table 1 and Fig. 3). A Friedman test (conducted with SPSS Statistics 26) did not reveal any significant differences between the experimental conditions ($\chi^2(3) = 4.79, p = 0.188, N = 28$).

Table 1: Descriptive statistics for Delta MSAQ scores across experimental conditions

| Condition | Mean | SD | Median | Min | Max |
|---------------|-------|-------|--------|--------|-------|
| Control | 14.68 | 18.67 | 6.60 | -2.78 | 68.75 |
| Visual cues | 13.72 | 15.08 | 8.68 | -0.69 | 51.39 |
| Dynamic seat | 11.76 | 15.86 | 7.64 | -13.19 | 61.11 |
| Reclined seat | 12.00 | 18.12 | 3.13 | -4.17 | 76.39 |

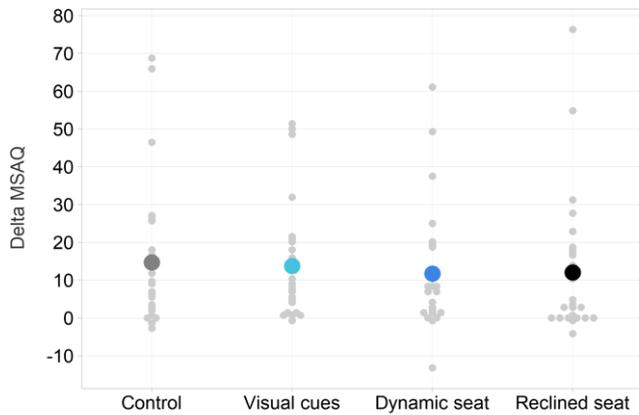


Fig 3: Distribution of individual scores for Delta MSAQ. Coloured circles indicate the mean value for each experimental condition.

3.2 Inter-individual differences

Figure 4 shows the effectiveness of all three countermeasure conditions in relation to the control condition by comparing the increase/decrease of the MSAQ scores on an individual level. This explorative analysis highlights the large inter-individual differences: Each countermeasure

| ID | Visual cues | Dynamic seat | Reclined seat |
|----|-------------|--------------|---------------|
| 1 | -2.08 | -4.86 | -5.56 |
| 2 | -0.69 | 0.00 | 0.00 |
| 3 | -6.25 | -5.56 | -18.06 |
| 4 | 4.17 | -9.72 | -13.89 |
| 5 | -17.36 | -16.67 | -34.72 |
| 6 | 10.42 | 13.89 | 6.25 |
| 7 | -16.67 | -6.94 | -9.72 |
| 8 | 6.25 | 4.17 | 4.17 |
| 9 | 2.08 | -2.78 | -6.94 |
| 10 | -1.39 | -0.69 | -2.08 |
| 11 | 2.78 | 0.00 | 0.00 |
| 12 | -4.17 | 6.94 | 4.86 |
| 13 | 9.03 | 13.19 | -2.78 |
| 14 | -11.81 | -38.89 | -2.78 |
| 15 | -8.33 | -24.31 | -8.33 |
| 16 | 20.14 | -12.50 | -11.81 |
| 17 | -22.92 | -14.58 | -13.19 |
| 18 | -2.08 | 5.56 | -2.78 |
| 19 | 1.39 | 2.08 | 27.78 |
| 20 | 3.47 | 0.69 | 0.69 |
| 21 | -17.36 | -31.25 | 7.64 |
| 22 | 6.94 | 4.86 | 1.39 |
| 23 | -5.56 | -6.25 | -6.25 |
| 24 | 10.42 | 0.69 | -6.94 |
| 25 | 3.47 | 14.58 | 8.33 |
| 26 | -0.69 | 15.97 | 7.64 |
| 27 | 6.25 | 0.69 | 0.00 |
| 28 | 3.47 | 9.72 | 2.08 |

Fig 4: Individual impact of countermeasures compared to the control condition (in terms of differences between Delta MSAQ scores).

seems to mitigate motion sickness for single participants (see negative values/green cells in Fig. 4), but not consistently across the test sample. Conversely, there is a clear indication that all countermeasures have the potential to increase motion sickness compared to the control condition (see positive values/red cells in Fig. 4). This result is also reflected in the verbal statements after the test drive.

3.3 Relationship between self-reported motion sickness susceptibility and motion sickness occurrence

Pearson correlation coefficients were computed to assess the linear relationship between the MSSQ Short scores and Delta MSAQ scores for each experimental condition:

- Control condition: $r(26) = .15, p = .460$
- Visual anticipatory cues: $r(26) = .38, p = .047$
- Dynamic seat adjustment: $r(26) = .30, p = .125$
- Reclined seating position: $r(26) = .19, p = .327$

4. Discussion and conclusion

This study was designed to (1) replicate existing findings for selected measures aiming to mitigate motion sickness in automated vehicles and (2) to compare their effectiveness in a controlled setting. Although great care has been taken to control for potential confounding factors (e.g., within subject design, same time of day, full permutation of trials, standardized instructions) the positive effects of the countermeasures found in related studies (see chapter 1) could not be confirmed under the circumstances of this study. Statistically, this can be explained by minor differences between means and especially by large differences between individual responses within the experimental conditions. But what are the determining factors that may explain the large variance within each experimental condition? Apart from the spread of self-reported susceptibility to motion sickness, the large variance within the experimental conditions may also result from intra-individual (day-to-day) variations in motion sickness susceptibility or from individually different reactions to the countermeasures presented in the experiment. In order to control for the day-to-day effect, several measurements for identical conditions would be needed. On the other hand, verbal feedback also indicates that differences in means could be enlarged by improving the implementation of countermeasures. Both aspects should be carefully considered for follow-up studies. After all, providing effective means to mitigate motion sickness is crucial to enable NDRTs in highly automated vehicles.

5. Acknowledgments

This work results from the joint project RUMBA and has been funded by the Federal Ministry for Economic Affairs and Climate Protection on a resolution of the German Bundestag (project number 19A20007A).

References

Bohrmann, D., & Bengler, K. (2020). Reclined posture for enabling autonomous driving. In T. Ahram, W. Karwowski, S. Pickl, R. Taiar (eds), *Human Systems Engineering and Design II. IHSED 2019. Advances in Intelligent Systems and Computing, vol 1026* In International Conference on Human Systems Engineering and Design: Future Trends and

Applications (pp. 169–175). Springer, Cham.
https://doi.org/10.1007/978-3-030-27928-8_26

Brietzke, A., Pham Xuan, R., Dettmann, A., & Bullinger, A. C. (2021). Influence of dynamic stimulation, visual perception and individual susceptibility to car sickness during controlled stop-and-go driving. *Forschung im Ingenieurwesen*, 85, 517–526.
<https://doi.org/10.1007/s10010-021-00441-6>

Donohew, B. E., & Griffin, M. J. (2009). Motion sickness with fully roll-compensated lateral oscillation: effect of oscillation frequency. *Aviation, Space, and Environmental Medicine*, 80(2), 94–101.
<https://doi.org/10.3357/ASEM.2345.2009>

Gianaros, P. J., Muth, E. R., Mordkoff, J. T., Levine, M. E., & Stern, R. M. (2001). A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, space, and environmental medicine*, 72(2), 115–119.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2910410/>

Golding J. F. (2006). Predicting individual differences in motion sickness susceptibility by questionnaire. *Personality and Individual differences*, 41(2), 237–248.
<https://doi.org/10.1016/j.paid.2006.01.012>

Karjanto, J., Yusof, N. M., Wang, C., Terken, J., Delbressine, F., & Rauterberg, M. (2018). The effect of peripheral visual feedforward system in enhancing situation awareness and mitigating motion sickness in fully automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 58, 678–692.
<https://doi.org/10.1016/j.trf.2018.06.046>

Kuiper, O. X., Bos, J. E., Diels, C., & Schmidt, E. A. (2020). Knowing what's coming: Anticipatory audio cues can mitigate motion sickness. *Applied Ergonomics*, 85, 103068.
<https://doi.org/10.1016/j.apergo.2020.103068>

Kuiper, O. X., Bos, J. E., & Diels, C. (2018). Looking forward: In-vehicle auxiliary display positioning affects carsickness. *Applied Ergonomics*, 68, 169–175.
<https://doi.org/10.1016/j.apergo.2017.11.002>

Reason, J. T. (1978). Motion Sickness Adaptation: A Neural Mismatch Model. *Journal of the Royal Society of Medicine*, 71(11), 819–829.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1436193/pdf/jrsocmed00291-0059.pdf>

SAE International. (2021). *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*.
https://www.sae.org/standards/content/j3016_202104/

Deriving Extended Keystroke Level Model Resumability Operators: An Occlusion Study

Heather Wilson, David R. Large*, Gary Burnett

Human Factors Research Group, University of Nottingham, Nottingham. UK

*David.R.Large@nottingham.ac.uk

Abstract: Theoretical techniques to model and predict drivers' visual behaviour during the execution of secondary in-vehicle tasks, such as the extended keystroke level model (eKLM), are predicated on perfect task resumability during the interleaving of attention, i.e., it is assumed that the secondary task can resume without penalty – irrespective of task characteristics – as soon as attention is redirected to it. In practice, this is unlikely to be the case. Moreover, it is reasonable to opine that resumability may improve over increasing numbers of glances. A formative occlusion study was devised to explore search-and-touch HMI interactions in which the number of glances and task complexity varied, with the aim of deriving new eKLM resumability operators.

1. Introduction

Predictive modelling is a valuable tool that can be used to determine the expected level of visual demand associated with secondary tasks in a vehicle without extensive user testing (see: Large et al., 2018). However, modelling techniques often fail to account for any changes in performance due to the interleaving of primary and secondary tasks, notably when the secondary task requires multiple off-road glances. In theory, predictive models often assume that the secondary task can resume without interruption as soon as the driver re-engages their visual attention and that each subsequent glance has the same net effect on resumability. However, it is reasonable to opine that, in practice, a driver may be better prepared to resume the secondary task in each subsequent glance due to cumulative increases in task familiarity and mental preparation.

2. Background

2.1 Keystroke Level Model

The Keystroke Level Model (KLM) (Card, Moran, Newell, 1980) aims to predict how long it will take an expert user to accomplish a routine task without errors using an interactive computer system.

KLM reduces the elements of interacting with an interface or system to individual operators, each representing a unique activity, such as mental preparation, moving, pointing, swiping, etc. Each operator has a defined time allowance and thus, complete task-time can be calculated by summing individual times in the sequence they must occur.

In an automotive context, KLM task-time predictions can be used to estimate the visual demand (i.e. total duration of off-road glances) associated with an in-vehicle human-machine interface (IV-HMI). However, this may be distributed over several glances, with the driver redirecting their attention to the road during the intervening time.

KLM predictions fail to adequately articulate this interleaved attention. Indeed, directly equating task-time with off-road glance time fails to consider activities that may be possible (or may continue or conclude) without vision. It also fails to recognise that a driver will need to refamiliarise themselves with the secondary task in each subsequent glance.

The extended-KLM (eKLM) (Pettitt, Burnett & Stevens, 2007) aims to address the first point by incorporating the underlying principles of the visual occlusion protocol in its predictions (ISO, 2017). The approach is predicated on three assumptions:

1. During periods of vision, the operator sequence can progress without interruption;
2. An operator that begins during a period of vision can continue into an occluded period *providing* it is not reliant on the provision of vision (e.g., moving hand towards HMI);
3. An operator can *only* begin in an occluded period when vision is not required at any point in its duration (e.g., a keystroke operation in which the finger has already been moved to its target).

The authors demonstrate significant correlations between observed and predicted results in a validation study (Pettitt, Burnett & Stevens, 2007). Nevertheless, the technique assumes that secondary task operators requiring vision can resume without penalty in each new glance (i.e. perfect resumability), and therefore fails to account for potential changes in behaviour and performance over multiple glances.

3. Method

3.1 Overview

An occlusion study was devised to explore point-and-touch HMI interactions, in which the number of glances was enforced (1, 2 or 3). Participants were always required to select the target in the final glance. For example, if provided with two glances, they were instructed to use the first glance to visually acquire the target and mentally prepare, and the second glance to complete the task; participants were only allowed to move their hand/arm during the final glance, and were told to do so only when they were confident that they could make the correct selection (thereby effectively segregating searching and pointing time). Task complexity (based on the structuring of the interface, number of elements and Fitts' index of difficulty (Fitts, 1954)) varied.

3.2 Participants and Experimental Setup

Twelve experienced drivers took part. Participants sat in the University of Nottingham Human Factors driving simulator and were presented an in-vehicle touchscreen HMI located in the centre console of the vehicle.

Wearing occlusion glasses, participants were provided with either one, two or three “glances” for each task (and were told which before each task). Each glance provided 1.5s of vision (equivalent to a 2.0s off-road glance, in line with the occlusion protocol (ISO, 2017)). Thus, the total shutter open times (TSOT) were 1.5, 3.0, or 4.5s, respectively.

Targets were either 6mm (ID=6.0) or 12mm (ID=5.0) wide, based on relevant literature (Jin, Plocher & Kiff, 2007; Sesto, Irwin, Chen, Chourasia & Wiegmann, 2012), and presented as structured or unstructured arrays of 1, 4, 9 or 25 similar targets, arranged in a uniform square. Indices of difficulty (IDs) were calculated based on the width of the target and the distance from the participant’s hand to the target when their hand was placed on the steering wheel at the “10 o’clock” position, in line with Fitts (1954).

3.3 Procedure

For each condition (defined by the permissible number of glances – one, two or three), participants were required to find and select the target containing a designated number, which was spoken aloud to them. Participants completed 12 tasks for each condition, and therefore completed 36 tasks in total.

Tasks were presented in a randomised order to counteract learning effects. Participants were observed via cameras located inside the driving simulator. Videos were subsequently analysed using Behavioural Observation Research Interactive Software (BORIS) (Friard & Gamba, 2016) to determine task-times and accuracy (i.e., task performance). In addition, *resumability time* was defined as the time it took participants to resume the task in the final glance (i.e., the time between the start of the final glance and the moment the participant’s hand left the steering wheel).

3.4 Hypotheses

It was predicted that *resumability time* would reduce with increasing number of glances and increase with increasing task complexity. In addition, task performance (i.e., accuracy of item selection) was expected to improve with increasing number of glances and decrease with increasing task complexity.

4. Results

4.1 Task Resumability

Results show that the resumability time decreased with increasing number of glances, as expected. A Pearson’s Correlation analysis showed that the resumability time for each task was negatively correlated to the number of glances, such that as the number of glances increased, participants took significantly less time to resume the task ($r=0.792$, $p<0.01$). A Linear regression equation was derived showing this relationship: $y = 0.77 - 0.24x$. There was no significant correlation between ID (i.e., index of difficulty) and resumability time. However, complexity in terms of array size (i.e., number of potential targets) was positively

correlated with resumability time ($r=0.318$, $p<0.05$), indicating that it took longer to resume the task when there were more potential targets to select.

4.2 Task Performance

Results show that task performance improved with increasing number of glances, as expected (Pearson’s correlation: $r=0.385$, $p<0.05$). In addition, a negative correlation was found between ID (i.e., task complexity) and accuracy, such that as ID increased, accuracy decreased ($r=-0.550$, $p<0.01$).

4.3 Derived Resumability Time Operators

NHTSA guidelines (NHTSA, 2013) stipulate that an IV-HMI should only permit tasks expected to have an average success rate of 85% or higher. Resumability time was therefore derived using only data where accuracy was 85% or higher. In practice, this equated to tasks with an ID of 5.0 or less (i.e., 12mm targets) (which also suggests that smaller targets would not be permissible on an IV-HMI).

Resumability time (‘operators’) differed based on secondary task characteristics (i.e., structuring and number of targets.) Mean values are summarised in Table 1. These should be applied in accordance with existing eKLM assumptions.

Table 1 Derived eKLM Resumability Operators

| Number of Glances | Resumability Operator |
|-------------------|-----------------------|
| 1 | .44s |
| 2 | .29s |
| 3 | .27s |

5. Discussion and Conclusion

The results indicate a cumulative effect, i.e., the time to resume the task reduced in each successive glance. This suggests that drivers built up and retained residual task familiarity and could therefore respond more quickly in later glances. In addition, the results suggest that complex search-and-touch IV-HMI interactions could potentially be completed with greater efficacy if searching and pointing were segregated, i.e., by encouraging drivers to use an initial glance/s to mentally prepare/visually acquire their target and an additional, final glance to make the selection.

It is noted that to extract the required data, the number of glances was enforced, and there was no provision to execute the task sooner, even if the participant was ready; we also imposed a strict >85% accuracy criterion when deriving operators. This may have artificially extended task-time in some situations, e.g., for less-complex tasks undertaken over multiple glances, and limited formative data.

Using the newly derived *resumability* operators, predictions made using eKLM are likely to be more accurate, although findings will need to be validated with on-road/simulator data in further work.

References

- Card, S. K., Moran, T. P., & Newell, A. (1980). The keystroke-level model for user performance time with interactive systems. *Communications of the ACM*, 23(7), 396-410.
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 381-391. <https://doi.org/10.1037/h0055392>
- Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in ecology and evolution*, 7(11), 1325-1330.
- Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 4(1), 11-26. <https://doi.org/10.1080/17470215208416600>
- ISO 16673:2017 (2017). Road vehicles — Ergonomic aspects of transport information and control systems — Occlusion method to assess visual demand due to the use of in-vehicle systems. Geneva, Switzerland: International Organization for Standardization (ISO).
- Jin, Z. X., Plocher, T., & Kiff, L. (2007, July). Touch screen user interfaces for older adults: button size and spacing. In *International Conference on Universal Access in Human-Computer Interaction* (pp. 933-941). Springer, Berlin, Heidelberg.
- Large, D. R., Burnett, G., Crundall, E., van Loon, E., Eren, A. L., & Skrypchuk, L. (2018). Developing predictive equations to model the visual demand of in-vehicle touchscreen HMIs. *International Journal of Human-Computer Interaction*, 34(1), 1-14.
- NHTSA. (2013). Visual-Manual NHTSA Driver Distraction Guidelines For In-Vehicle Electronic Devices. NHTSA-2010-005
- Pettitt, M., Burnett, G., & Stevens, A. (2007). An extended keystroke level model (KLM) for predicting the visual demand of in-vehicle information systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1515-1524).
- Sesto, M. E., Irwin, C. B., Chen, K. B., Chourasia, A. O., & Wiegmann, D. A. (2012). Effect of touch screen button size and spacing on touch characteristics of users with and without disabilities. *Human Factors*, 54(3), 425-436.

Biomotion Triggering Driver (In)Attention for Cyclists at Night

Paul Hemeren^{1*}, Mikael Johannesson¹, Mikael Lebram¹, Fredrik Eriksson¹

¹ School of Informatics, University of Skövde, Box 408, 541 28 Skövde, Sweden

*paul.hemeren@his.se

Abstract: This research demonstrates the driver attention of cyclists wearing different patterns of reflectors in twelve different visual traffic conditions that occur in an existing city environment at night. A driving simulator was used to include more naturalistic testing conditions and to achieve a relatively high level of experimental control. The accuracy and distance at which drivers would detect cyclists was measured. Cyclists were dressed in three different patterns of reflective clothing, i.e., biomotion, standard vest, and no reflective material at all, which is the minimum legal requirement. A 4.6 kilometer long route which included central areas of the city of Skövde was selected for the video recording of cyclists dressed in the different reflective clothing conditions. The participants were instructed to drive the car and maintain a speed of 50 km/h as the primary task and to toot the horn if they saw a cyclist. The major overall result was the significantly greater distance at which drivers detected cyclists with biomotion reflective clothing than the standard vest and no reflective material. Whereas the differences between standard reflective vests and no reflective material were not significantly different in many of the traffic situations. These results demonstrate that driver attention is triggered in a secondary task for cyclist biomotion at distances that provide more time to care for cyclist safety.

1. Introduction

The use of visual aids can increase the ability of drivers to detect cyclists at night and reduce the seriousness of injuries if a crash occurs (Kwan & Mapstone, 2004; Wood, 2020). The use of reflectors placed on critical parts of the human body have been shown to increase cyclist conspicuity at night (Wood, Tyrell, Marszalek, et al., 2011). Drivers detect cyclists with reflective clothing that enhances the movement of the human body (biomotion) at considerably longer distances than a reflective vest, which is a very often used piece of clothing by cyclists who want to be detected in darkness. Research (Wood, Tyrell, Lacherez & Black, 2017) has also shown that driver eye movements are quicker to fixate on cyclists who are wearing the biomotion reflector clothing than the reflective vest.

The effect of the biomotion-patterned reflector placement stems from the sensitivity of human vision to the movement of other humans. This has been demonstrated in much research, which was promulgated by Gunnar Johannson (Johannson, 1973). This sensitivity can be exploited by placing reflective material on the joints of the human body so that visual conspicuity is increased for cyclists. Edwaard, et al., (2020) showed clear results of the conspicuity benefits of biological motion when cyclists were pedalling during the daytime.

This research project addresses the effectiveness of cyclists wearing different patterns of reflectors on the distance it takes drivers to detect cyclists in different visual conditions that occur in an existing city environment. As a complement to our previous research (Hemeren, et al., 2014) we aimed to determine the distance at which drivers would detect cyclists dressed in different patterns of reflective clothing, i.e., biomotion, standard vest, and no reflective material at all on the cyclist, which is the minimum legal requirement. This distance effect is also addressed in different traffic contexts that might interact with reflective material.

2. Method

2.1 Participants

Twenty-four participants (19 males, mean age 29) were recruited from the student population at the University of Skövde and from the circle of acquaintances of the experimenters. All participants had a valid driving license, but driving experience/frequency varied, from once a month to daily. All participants signed consent forms, and the experiment was conducted according to Swedish law and ethical guidelines.

2.2 Design and Procedure

Three clothing patterns were used (Fig. 1): biomotion, vest and the minimum legal requirement (legal), in which no reflector material was worn by the cyclists. The reflective material used in the biomotion and reflective vest conditions was 3M™ Scotchlite™ Reflective Material 8910 Silver Fabric. The reflective material on the reflective vest was replaced with the same reflective material used for the biomotion clothing. Importantly for comparison, the amount of visible reflective material in these two conditions was approximately the same.

Twelve positions were also selected along the route for the placement of cyclists. The positions were selected to include well-lit areas with street lights and other light sources for commercial areas, no street lights and positions where the visibility of cyclists was partially obstructed by making a turn or by bushes along the road. The bicycles were placed in stationary training stands so that distance measures could be reliably made while cyclists pedalled. Each clothing condition was video recorded at each position for a total of thirty-six conditions. The spacing of the positions along the route was unevenly spaced in order to reduce expectancy effects about the presence of a cyclist.

In the driving simulator, the instructions were to drive the car and maintain a speed of 50 km/h as the primary

task. The gas pedal of the car was used to increase the speed of the recorded film to give the impression of acceleration. The secondary task was to honk the horn when a cyclist was detected.

Each participant drove the route two times, which resulted in viewing each clothing condition six times for a total of 18 observations per participant. Since it was not practically feasible for each participant to view each clothing condition at each position, because of likely practice effects, the conditions were evenly divided among two groups of participants.

The pattern of results across the different positions/places shows that the biomotion pattern of reflector placement ($\bar{x} = 60.28$, $SEM=1.27$) is detected at much greater distances than legal ($\bar{x} = 31.67$, $SEM=1.32$) or vest ($\bar{x} = 33.56$, $SEM=1.32$). The exceptions to this occur when the visibility of cyclists is obstructed. The main effect of reflector placement was significant, $F(2,18) = 16.91$, $\eta^2 = 0.65$, $p < 0.001$. Pairwise comparisons between the conditions for reflector placement show significant differences between the legal and biomotion conditions, $t(227) = 10.46$, $p < 0.001$, and between the vest and biomotion conditions, $t(227) = 9.62$, $p < 0.001$. The main effect of position was also significant, $F(9,18) = 2.69$, $\eta^2 = 0.57$, $p < 0.001$. The interaction effect showed that the



Fig. 1. Reflective clothing patterns from left to right: biomotion, vest and legal minimum.

3. Results

There were twelve possible correctly detected cyclists for each of the thirty-six conditions due to the necessary balancing of conditions and to reduce practice effects. The detection results for each condition showed that drivers were generally accurate at detecting cyclists with the exception of two positions (9 and 11) along the route where detection accuracy for the legal condition (no reflective material) was 0 and 3 out of twelve. The detection accuracy for the vest condition for these two positions was also low, 4 and 6 respectively, while the biomotion condition resulted in perfect detection in all twelve trials.

The data for the distance at which cyclists were detected were entered into a between-groups ANOVA with position as a random factor. The means for detection distance as a function of clothing condition and route position are presented in Figure 2, which shows in positions 1, 3, 4, 5, 6, 7, 8, and 10 that the biomotion patterns triggered detection at much greater distances than the legal and vest conditions. In these positions, the street paths were fairly straight and could give rise to early detection. For positions 2 and 12, there were curves where there was no sufficient distance between the car and the bicycle in order to detect biomotion patterns at a greater distance.

effect of the different reflector conditions varied as a function of position along the route, $F(18, 309) = 9.16$, $\eta^2 = 0.35$, $p < 0.001$.

4. Conclusions

The reflective biomotion clothing that triggers attention is superior to the vest and the legal minimum at almost twice the average distance. It is also important to see that this effect, however, varies as a function of the different night time visual traffic situations. A further more surprising result here is the lack of any significant average difference between the vest and legal minimum conditions. This indicates that cyclists should wear reflective clothing that can trigger driver attention by the placing reflective material on the joints of the human body. Driver distraction in areas with many different sources of light is also significantly reduced when cyclist biomotion is reflected. An application of these results could be developed for computational models that are used to detect cyclists, not only at night but during the daytime as well.

5. Acknowledgments

This project is a result of research funding from the Swedish Transport Administration and from the University of Skövde.

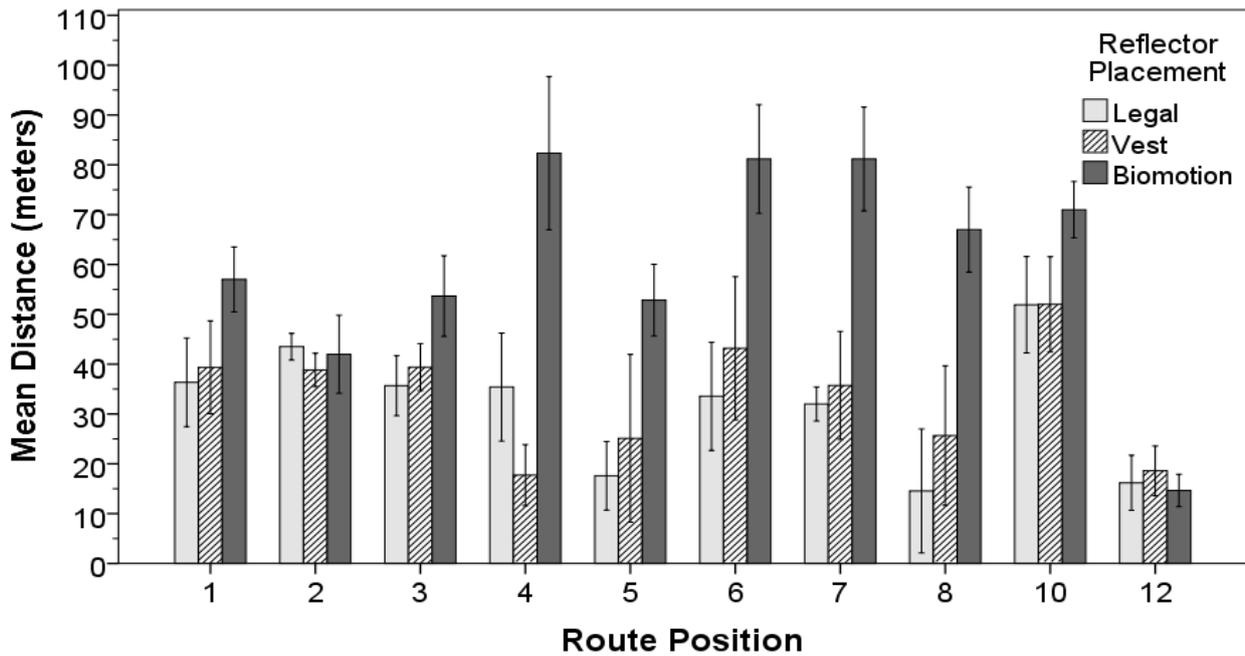


Fig. 2. Distance detection in meters as function of reflector condition and place on route. Positions 9 and 11 are not included due to the very low detection accuracy for the legal and vest conditions.

References

- Edewaard, D. E., Fekety, D. K., Szubski, E. C., & Tyrrell, R. A. (2020). Highlighting bicyclist biological motion enhances their conspicuity in daylight. *Accident Analysis & Prevention*, 142, 105575.
- Hemeren, P. E., Johannesson, M., Lebram, M., Eriksson, F., Ekman, K., & Veto, P. (2014). The use of visual cues to determine the intent of cyclists in traffic. In *2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)* (pp. 47-51). IEEE.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & psychophysics*, 14(2), 201-211.
- Kwan, I. & Mapstone, J. (2004). Visibility aids for pedestrians and cyclists: a systematic review of randomised controlled trials. *Accident Analysis and Prevention*, 36, pp. 305-312.
- Wood, J. M. (2020). Nighttime driving: visual, lighting and visibility challenges. *Ophthalmic and physiological optics*, 40(2), 187-201.
- Wood, J. M., Tyrrell, R.A, Marszalek, R. P., Lacherez, P. F., Carberry, T. P. & Chu, B. S. (2011). Using reflective clothing to enhance the conspicuity of bicyclists at night. *Accident Analysis and Prevention*, 45, pp. 726-730.
- Wood, J. M., Tyrrell, R.A, Lacherez, P. & Black, A. A. (2017). Night-time pedestrian conspicuity: effects of clothing on drivers' eye movements. *Ophthalmic & Physiological Optics*, 37, pp. 184-190.

Effect of agency on the occurrence of car sickness: incidence for autonomous driving

Eléonore Henry^{1,2*}, Clément Bougard^{1,2}, Christophe Bourdin², Lionel Bringoux²

¹ Stellantis, Centre Technique de Vélizy, Vélizy-Villacoublay, France

² Aix Marseille Univ, CNRS, ISM, Marseille, France

eleonore.henry@stellantis.com

Abstract: With the development of autonomous vehicles, drivers will become passengers. This evolution leads to a loss of agency that could affect the perceptual states of vehicle occupants and thus, induce car sickness for drivers. This study sought to identify the influence of agency levels on the severity of car sickness. 16 healthy subjects participated as front passengers/drivers (dual control vehicle) in a slalom session with oscillating lateral movements at 0.2Hz. Four different conditions were proposed: (1) passengers, (2) passengers with hands on the steering wheel (not driving), (3) drivers and (4) drivers with interventions of a simulated ADAS. Every test session comprised a baseline (5min), slalom (20min), recovery (5min) periods. For each period, participant level of car sickness was recorded. Before and after each test session participants completed the MSAQ. Finally, after each test session, the level of perceived control (LPC) and felt at ease with vehicle path were assessed. Our results reveal that car sickness severity is higher for passengers than drivers. Also, driving with ADAS interventions disrupts LPC of vehicle path, limiting car sickness recovery. In addition, passengers getting proprioceptive information by having their hands on the steering wheel exhibited increased LPC and reduced car sickness severity. For the first time, our results demonstrated, the importance of developing solutions allowing for better human-vehicle coupling, to limit the severity of car sickness due to autonomous driving, notably by improving the feeling of agency.

1. Introduction

Cars, the most common form of transportation, could induce a specific type of motion sickness (MS) referred to as car sickness (CS) (Murdin et al., 2011). Nowadays, passengers are the most likely to develop CS symptoms while drivers remain almost unaffected (Rolnick and Lubow, 1991). However, with the current development of autonomous driving functions and autonomous vehicles, drivers will become passengers in their own vehicle (Wada, 2016). Consequently, the number of car occupants affected by CS could increase significantly (Kuiper et al., 2018).

This difference in susceptibility to CS between passengers and drivers depends on several factors: control, activity, anticipation (Wada, 2016). These factors partly refer to the notion of agency, which defines the level of intentional control over an action (Haggard and Chambon, 2012). Being agent in a situation can help mobilize attentional mechanisms and improve sensorimotor integration resulting in better performance in human machine interactions (Berberian et al., 2012). This is supported by recent theories which suggest that a match between internal expectations of motor commands and current perceptual estimates issued from sensory-feedback allows for a better sense of agency (Haggard, 2017) and thus could limit the occurrence of MS (Bos et al, 2010).

In this way, we hypothesize that the level of agency may play a major role in the way driving situations may impact perceptually-related states of operators differently engaged in driving task. The aim of the study is then to evaluate for the first time the

influence of agency on CS symptoms severity in real driving conditions.

2. Materials and Methods

2.1 Participants

Sixteen healthy participants with high susceptibility to MS (8 women, 8 men, age: 40.4 ± 8 years) volunteered to take part in this study.

2.2 Experimental set-up

Test sessions were conducted in a closed area. Experimental road consisted of two straight segments with radius turning zones at both ends. Twelve pylons were located 20m apart along both straight segments. This configuration and car speed (35 km/h) ensured lateral movements of 0.2Hz (slaloms), recognized as CS-inducing frequency (Henry et al., 2022). Vehicle used was a dual control vehicle (C4 Picasso) to enable the simulation of autonomous driving functions with a professional driver in front passenger seat (hidden by a partition). Participants were seated in the driver position. Four conditions were developed to modify the level of agency: Passengers (P - very low level), Augmented Passengers (AP - low level), Disturbed Drivers (DD - high level) and Drivers (D - very high level). For P and AP conditions, driving task was performed by a professional driver and participants thought to be in an automated vehicle. Participants had their hands on their knees (P) or on the steering wheel (AP). In DD and D conditions, participants were required to perform themselves the slalom driving task. For DD condition, several unexpected interventions on the steering wheel mimicking

interventions of an advanced driver assistance system (ADAS) were produced by the professional driver. All participants took part in the four conditions in a random order.

2.3 Procedure and data acquisition

Test sessions comprised a baseline period in static conditions (5min), a slalom period in dynamic conditions to induce CS (20 min) and a static recovery period (5min). For each slalom, CS ratings (CSR) was recorded using a 5 points continuous likert-scale (0: no symptoms – 4: moderate nausea) (Griffin and Newman, 2004). If participants reached level 4, slalom session was stopped. For baseline period, only one value was recorded and for recuperation period, participants were instructed to indicate their CS level every minute (5 values). Before and after each test session, participants completed a MS assessment questionnaire (MSAQ) used to evaluate four dimensions of MS, which were defined as gastrointestinal, central, peripheral, and sopite-related (Gianaros et al, 2003). Finally, to infer the level of perceived agency in the 4 conditions, we asked participants after each test session, about their level of perceived control (LPC) of vehicle path (0: no control - 10: total control) and whether they felt at ease (FE) with vehicle path (0: not at all - 10: totally).

2.4 Statistical Analysis

Three dependent variables were analyzed: (i) level of perceived agency (LPC and FE), (ii) delta value (% of variation between pre-post) for MSAQ scores (Total, Gastrointestinal, Central, Peripheral, and Sopite) and (iii) CSR. First two variables were analyzed using a 4-level (conditions: P, AP, DD and D) Friedman ANOVA for each element. When significant differences were observed ($p < 0.05$),

Wilcoxon tests were performed. For CSR, 9 periods were selected: 4 for slalom (slalom with the highest score (S_{max}) and first (S_{start}), middle (S_{mid}), final slalom (S_{stop})) and 5 for recuperation periods (1-5min, R_{1-5}). CSR were analyzed using a 4-level (conditions: P, AP, DD and D) \times 9 periods repeated measures ANOVA. When significant differences were observed ($p < 0.05$), LSD post-hoc analysis was performed. Statistical analyses were achieved using Statistica software® v.10 (Statsoft Inc, France).

3. Main results

3.1 Level of perceived agency

A significant effect of the ‘conditions’ was observed on the LPC of vehicle path. Wilcoxon tests revealed that LPC was significantly higher for D compared to DD, P and AP conditions. In addition, LPC was higher in DD than in P condition. Finally, LPC in AP was higher than in P condition (Figure 1).

A significant effect of the ‘conditions’ was observed on the FE with the vehicle path (Figure 1). Wilcoxon tests revealed that FE was significantly higher for D compared to P condition, and for AP compared to P condition (trend observed).

3.2 MSAQ Scores

For the MSAQ scores, a significant effect of the ‘conditions’ was observed only in drivers versus passengers conditions. Wilcoxon tests revealed that Total scores was significantly higher for P and AP conditions compared to D and DD conditions (Table 1). For Gastrointestinal scores, P condition increased significantly compared to D and DD conditions. For Central scores, P condition showed higher values than DD condition. Finally, for Peripheral scores, passengers conditions increased significantly compared to drivers conditions.

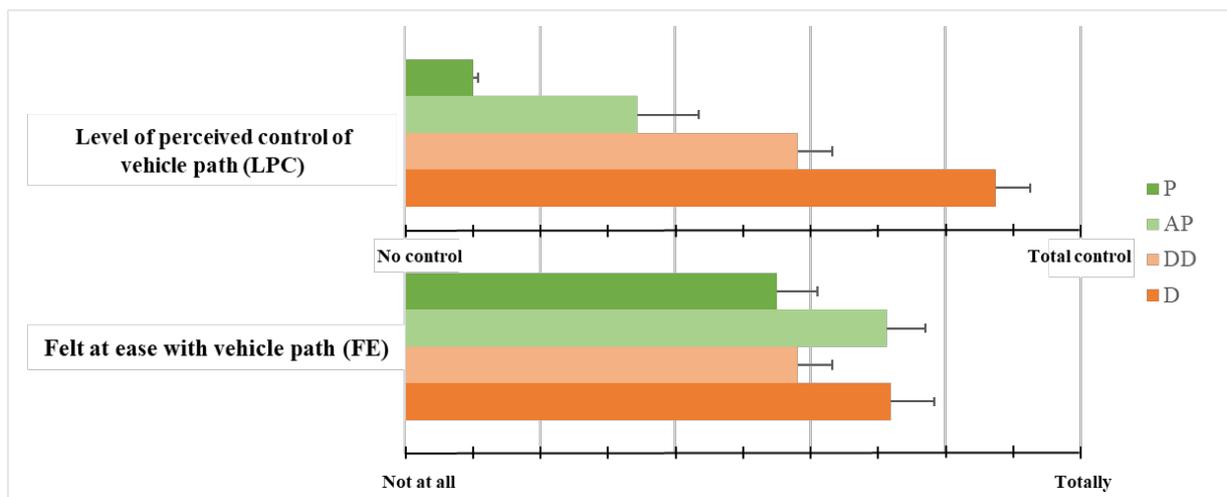


Figure 1. Level of perceived agency observed for each condition (mean \pm SEM; n = 16)

Table 1 Delta value for MSAQ scores observed for each condition (mean % \pm SEM; n = 16)

| | P | AP | DD | D |
|------------------------------|----------------------|----------------------|----------------------|---------------------|
| MSAQ Total | + 114,9 % \pm 28,7 | + 127,9 % \pm 32,0 | + 69,5 % \pm 17,4 | + 71,3 % \pm 17,8 |
| MSAQ Gastrointestinal | + 145,6 % \pm 47,1 | + 167,6 % \pm 34,2 | + 106,4 % \pm 30,2 | + 98,8 % \pm 35,2 |
| MSAQ Central | + 94,2 % \pm 32,7 | + 104,3 % \pm 26,1 | + 59,6 % \pm 26,2 | + 58,5 % \pm 19,2 |
| MSAQ Pheripheral | + 151,6 % \pm 48,1 | + 186,7 % \pm 50,8 | + 78,6 % \pm 32,0 | + 80,8 % \pm 34,7 |
| MSAQ Sopite | + 114,7 % \pm 32,5 | + 117,1 % \pm 23,1 | + 68,5 % \pm 20,4 | + 68,3 % \pm 19,9 |

3.3 Car-sickness ratings

A significant effect of the ‘conditions’ was observed on CSR measured by S_{Max} (Figure 2A). Post-hoc analyses indicated higher values for P and AP than D and DD conditions, P also differed from AP. This effect of ‘conditions’ on CSR is observed for all periods, except for S1, which is used as a

reference (Figure 2B). From S_{stop} , ratings were significantly higher for P condition than for AP, D and DD conditions. Finally, stopping the slalom induced a significant decrease during the recovery period for all conditions ($R_5 \neq R_1$). However, ratings returned to S1 level ($R_5 = S_1$) for D and AP conditions only.

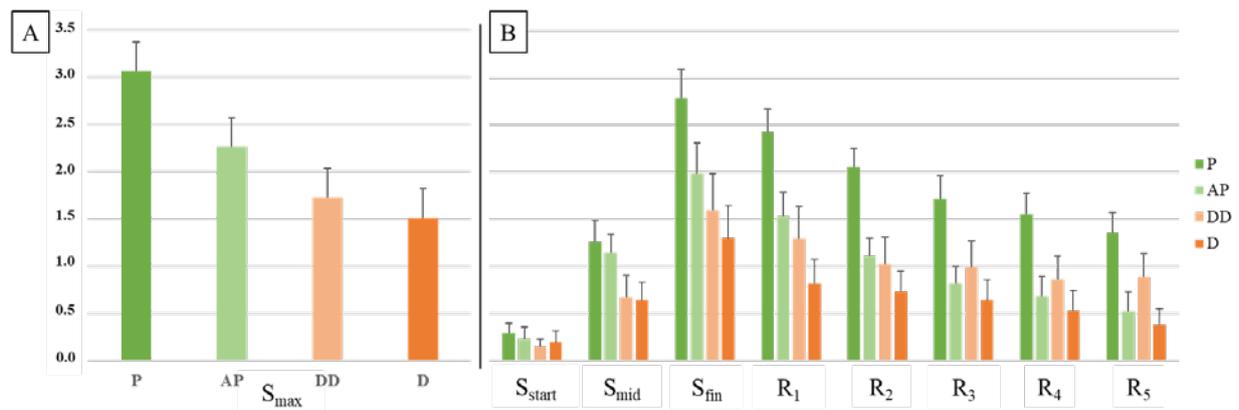


Figure 2 Level of CSR observed for each condition (A) for S_{max} (B) for slalom and recuperation periods (mean \pm SEM; n = 16).

4. Discussion

Our results show that LPC obtained in the driver conditions are higher than in the passenger conditions. For the CSR, an opposite effect was observed, all ratings obtained in the passenger conditions are higher than in the driver conditions. This is also supported by our MSAQ scores, in particular for the gastrointestinal dimension of CS. Our results confirm that drivers are less susceptible to CS than passengers (Dam et al., 2021). This supports the idea that improving the sense of control leads to a decrease in MS (Levine et al., 2014). Perceived control has been shown to inform about the sense of agency (David, Newen and Vogeley, 2008). It seems that the greater the sense of agency, the lower CSR.

More precisely, the LPC and the FE was higher in AP than in P condition. Placing passively the hands on the steering wheel during the driving task yielded an improved LPC and FE, and reduced CSR. This condition gives additional proprioceptive

information of the vehicle path compared to a classic passenger condition. It seems then that improving sensorimotor integration (Dong et al., 2011) could allow for a better sense of agency (Myers, Mock and Golob, 2020), limit the severity of CSR and allow CS recovery.

In contrast, the LPC and the FE was lower for DD compared to D condition. Driving with ADAS interventions disrupts the LPC and FE, and limits CS recovery. It seems that the sense of agency is optimal when voluntary actions match sensory-related outcomes. However, when they mismatch, errors in predicted states occur (Haggard, 2017). Larger and repeated errors of prediction during the slalom period could explain why participants do not recover entirely at the end of the test session. Alternatively, with ADAS interventions, attentional mechanisms were mobilized that could prevent the occurrence of CS until the stimuli stopped.

5. Conclusion

For the first time, our results show that the use of solutions restoring the sense of agency (by a better human-vehicle coupling) may help limit the occurrence of CS due to driving automation.

Acknowledgments

This study was part of the OpenLab agreement “Automotive Motion Lab” between Stellantis and Aix-Marseille University and CNRS. We thank to Rabah Sadoun and Samir Bouaziz (from Laboratory SATIE – Paris Saclay) the preparation of the acquisition material for this experiment. Finally, we thank to participants for their availability.

Abbreviations

(ADAS), advanced driver assistance system; (AP), Augmented Passengers; (CS), car sickness ; (CSR), car sickness ratings; (D), Drivers; (DD), Disturbed Drivers; (FE), Felt at ease ; (LPC), level of perceived control; (MS), motion sickness; (MSAQ), motion sickness assessment questionnaire; (P), Passengers; (R₁), rest for 1min; (R₂), 2min; (R₃), 3min; (R₄), 4min; (R₅), 5min; (S_{start}), first slalom; (S_{mid}), middle slalom; (S_{stop}), final slalom; (S_{max}), slalom with the highest score.

References

Berberian, B., Sarrazin, J. C., Le Blaye, P., & Haggard, P. (2012). Automation technology and sense of control: a window on human agency. *PloS one*, 7(3), e34075.

Bos, J. E., de Vries, S. C., van Emmerik, M. L., & Groen, E. L. (2010). The effect of internal and external fields of view on visually induced motion sickness. *Applied ergonomics*, 41(4), 516-521.

Dam, A., & Jeon, M. (2021, September). A Review of Motion Sickness in Automated Vehicles. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 39-48).

David, N., Newen, A., & Vogeley, K. (2008). The “sense of agency” and its underlying cognitive and neural mechanisms. *Consciousness and cognition*, 17(2), 523-534.

Dong, X., Yoshida, K., & Stoffregen, T. A. (2011). Control of a virtual vehicle influences postural activity and motion sickness. *Journal of Experimental Psychology: Applied*, 17(2), 128.

Gianaros, P. J., Muth, E. R., Mordkoff, J. T., Levine, M. E., & Stern, R. M. (2001). A Questionnaire for the Assessment of the Multiple Dimensions of Motion Sickness. *Aviation Space and Environmental Medicine*, 72(2), 115–119.

Griffin, M. J., & Newman, M. M. (2004). Visual field effects on motion sickness in cars. *Aviation, space, and environmental medicine*, 75(9), 739-748.

Levine, M. E., Stern, R. M., & Koch, K. L. (2014). Enhanced perceptions of control and predictability reduce motion-induced nausea and

gastric dysrhythmia. *Experimental brain research*, 232(8), 2675-2684.

Haggard, P., & Chambon, V. (2012). Sense of agency. *Current biology*, 22(10), R390-R392.

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, 18(4), 196-207.

Henry, E. H., Bougard, C., Bourdin, C., & Bringoux, L. (2021). Changes in Electroencephalography Activity of Sensory Areas Linked to Car Sickness in Real Driving Conditions. *Frontiers in Human Neuroscience*, 15.

Kuiper, O. X., Bos, J. E., and Diels, C. (2018). Looking forward: in-vehicle auxiliary display positioning affects carsickness. *Appl. Ergon.* 68, 169–175. doi: 10.1016/j.apergo.2017.11.002

Murkin, L., Golding, J., and Bronstein, A. (2011). Managing motion sickness. *BMJ* 343:d7430. doi: 10.1136/bmj.d7430

Myers, J.C., Mock, J.R. and Golob, E.J. (2020) ‘Sensorimotor Integration Can Enhance Auditory Perception’, *Scientific Reports*, 10(1), p. 1496. doi:10.1038/s41598-020-58447-z.

Rolnick, A., & Lubow, R. E. (1991). Why is the driver rarely motion sick? The role of controllability in motion sickness. *Ergonomics*, 34(7), 867-879.

Wada, T. (2016, September). Motion sickness in automated vehicles. In *Advanced Vehicle Control: Proceedings of the 13th International Symposium on Advanced Vehicle Control (AVEC’16)* (pp. 169-174).

Ocular activity reflecting mental models' dynamics is influenced by autonomous driving duration.

Arthur Portron^{1*}, Gaëtan Perrotte^{1,2}, Christophe Bourdin¹, Clément Bougard^{1,2}, Jean-Louis Vercher¹

¹ Aix Marseille Univ, CNRS, ISM, Marseille, France

² Stellantis - Centre Technique de Vélizy, Vélizy-Villacoublay, France

*arthur.portron@univ-amu.fr

Abstract: Level 3 semi-autonomous function should allow the driver to become a passenger and engage in a secondary task while the function is activated, and then return to the driving task when the system requests it. The presence of a non-driving related task, as well as high level of trust in the system, is known to impact the supervision of the driving environment by the driver, which is a major safety issue. Thus, while high knowledge of the system causes a decrease in eye activity in the driving environment, we know less about how this activity evolves as the driver's mental model may evolve. We propose then to investigate how driver's mental model dynamic is influenced by autonomous driving duration through the analysis of gaze activity. In a driving simulator, 51 participants went through conditional automated driving scenarios for two sequences, lasting either 2x5 minutes, 2x15 minutes, 2x45 minutes or 2x60 minutes (namely conditions C05, C15, C45 and C60), with a non-driving related task (watch a movie). Results showed that during a drive session, the more time passes, the less drivers monitored the driving environment, and consequently drivers who had experienced short durations were more likely to sustain their monitoring activity on the driving environment than drivers in the long duration groups. This result also suggests that in the second session participants try to anticipate the appearance of takeover request and behave according to what they were exposed to in the first sequence.

1. Introduction

During Level 3 of semi-autonomous function, the driver is allowed to be out-of-the-loop (OOTL) as he/she can be involved in a non-driving related task but must be able to resume control of the vehicle when the system requests it (via a takeover request or TOR). Thus semi-autonomous driving functions restrict the role of the driver to a supervisor and during most of the travel time to a passenger. These changes imply that the control loops required in manual driving are no longer activated with the same dynamic and efficiency. Control loops theories come from aeronautic research about pilot interaction with aircraft and were formalized first by Kaber and Endsley (1997), then lately updated by Merat et al. (2019). According to this latter, manual driving means that the driver is "in the loop" i.e. in the physical loop of control of the vehicle (trajectory and speed control) while the cognitive loops include the processes related to environment supervision and navigation.

To ensure safety, further studies are needed to evaluate how human factors influence the OOTL state, the way the driver behaves during the non-driving period, and what are the driver-automata interaction dynamics. However, these questions are difficult to address because this function is not still available and mental model assessment methods are indirect. According to Carroll & Olson (1988), mental models are "elaborate structures reflecting the user's understanding of what the system contains, how it works, and why it works that way". A lot of studies on trust and acceptance established that a better knowledge of the system, brought by explicit learning, increases these two dimensions (Beggiato & Krems,

2012; Metz et al, 2021) but also that a really accurate mental model of the system can lead to fast OOTL states due to complacency (Endsley, 1995, Parasuraman et al, 2000, Bahner et al, 2008). In level 3, the non-driving related task implies an OOTL state which could be countered by the voluntary decision by the driver to monitor the environment (as he/she knows that at some time he/she must regain control of the car). Link between mental model and ocular activity is established from aeronautic studies with pilots who showed specific oculomotor activity as they have extremely accurate mental model of their operation system (Lounis et al., 2021).

A self-initiated strategy to monitor driving environment in level 3 autonomous context matches with a mental model involving the necessity to takeover the system at some point and the need for information to ensure a safe takeover. This matches with an "On the loop state" as the driver activates one particular sub-task of the driving activity such as hazard perception and perception of anticipatory cues (Stanton et al, 2001). If it is known that factors as NDRTs compete with gaze activity on driving environment relative to their cognitive demand (Feldhütter et al, 2017; Du et al, 2020), less is known on how oculomotor activity could be modulated as the expectation of the driver evolve through time.

From a dataset collected in previous experiment (Portron et al, submitted), we propose to investigate how gaze activity during the autonomous period is influenced by experience in a two successive exposures design with a passive non-driving related task (no performance expected within the NDRT) in short and long duration conditions.

2. Hypothesis

We expect that the experience of different durations of autonomous conditions may shape the gaze activity as the participants build a mental model based on near past experience. Furthermore, we expect that participants in short duration of autonomous conditions tend to monitor the driving environment more actively than participants in longer duration conditions, and this more especially during the second exposure.

3. Material and methods

3.1 Driving procedure

In a static driving simulator located at the Center for Virtual Reality of the Mediterranean facility, 51 participants went through conditional automated driving scenarios for two sequences (sequence A and B) of same duration (either 5 minutes, 15 minutes, 45 minutes or 60 minutes, namely conditions C05 (n= 13), C15 (n = 13), C45 (n = 13) and C60 (n = 12)) with a 10 minutes break between the two sequences. In each sequence, participants had to react to a take-over request and avoid a simulated road event in front of

Table 1. Summary of total visit duration (mean (\pm sd)) for groups duration and sequences with p-value from ANOVA and effect size (η^2).

| Mean Duration Visit: TOI Post_MtoA | | | | | | |
|------------------------------------|---------------|---------------|---------------|---------------|-------|----------|
| Time | C05 | C15 | C45 | C60 | p | η^2 |
| Sequence A | 18.47 (13.68) | 25.25 (12.91) | 13.05 (9.56) | 17.24 (14.72) | 0.136 | - |
| Sequence B | 16.46 (13.50) | 17.28 (10.50) | 11.77 (13.66) | 11.56 (8.28) | 0.491 | - |
| Mean Duration Visit: TOI Pre_TOR | | | | | | |
| Time | C05 | C15 | C45 | C60 | p | η^2 |
| Sequence A | 12.94 (11.72) | 14.5 (12.84) | 6.12 (11.61) | 8.70 (9.26) | 0.252 | - |
| Sequence B | 19.44 (15.99) | 9.94 (10.10) | 4.27 (5.05) | 2.04 (2.53) | 0.000 | 0.33 |

Declaration of Helsinki and all participants received detailed information about the study and gave informed consent. The protocol was approved by a Bio-ethical committee (CERSTAPS, IRB00012476-2020-15-07-63).

3.2 Data acquisition and processing

Ocular activity was recorded by using a Tobii® Pro Glasses 2 system at 50 Hz. We defined specific time of interest (see Fig.1, mostly based of the design from Gold et al., 2017) and one specific area of interest: for the driving environment (AOI_DE, see Fig 1A).

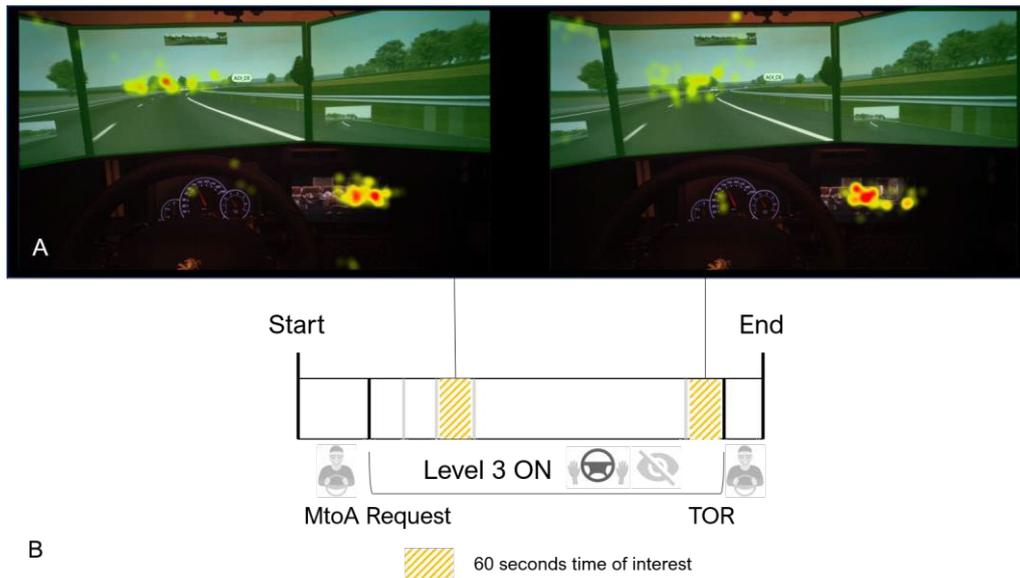


Figure 1 : A: Example of a heatmap for each time of interest (TOI) during the first sequence for one participant. Heatmap illustrates fixation time and distribution of one participant in the first sequence of the visual field with a maximum scale value (red area) for a fixation duration of 3 seconds. Green area on each snapshot illustrate the area of interest. B: Schema of the times of interest. The first TOI (post_MtoA) of 60 seconds is defined from the 3rd minute after the activation of the autonomous function. The second TOI (preTOR) is defined from the last 60 seconds before the takeover request.

the drive. During the autonomous driving period, participants were allowed to watch a movie ("Aquaman", 2018, WarnerBros ©) on a screen located on the right of the steering wheel and they were informed that the TOR notification and so the road-event could happen at any time. This study was performed in accordance with the principles of the

3.3 Dependent variables

Ocular activity was assessed through the mean total duration of visits on the AOI, which corresponds to the mean of cumulated fixation times over a 60 seconds period. This feature was calculated for two different times of interest. According to the design of

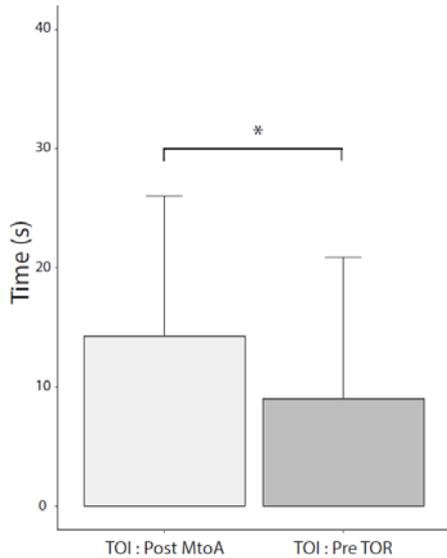


Figure 2. Barplots illustrate mean total duration of visits on driving environment for the two times of interest in sequence B. Vertical black lines illustrate standard deviation for each mean value. * for $p < 0.05$.

the experiment, the first times of interest (TOI) took into account the ocular activity observed during 60 seconds from the 3rd minute after the activation of the autonomous function (Post_MtoA). The second TOI refers to the 60 seconds just before the TOR notification (Pre_TOR), i.e. after 4 minutes (C05 group ;), 14 minutes (C15 group), 44 minutes (C45 group) and 59 minutes (C60 group), respectively.

3.4 Statistical analysis

The mean total duration of visits on driving environment was analysed by a one-way ANOVA with groups of duration as factor and a paired one-way ANOVA for the TOIs as factors for each sequence. If significant differences were observed ($p < 0.05$), a post-hoc Tukey HSD analysis was conducted. Effect size is expressed with eta squared and its interpretation is based on Cohen's rules (1992). Mean values are expressed with their standard deviation as follows (Mean (\pm SD)).

4. Results

4.1 Mean total duration spent on driving environment.

4.1.1 Sequence A

The ANOVA analysis revealed a significant difference for the TOIs ($F(1,98) = 10.58$, $p < 0.01$, $\eta^2 = 0.10$ (small)). We observed a greater time spent on the driving environment after the activation of the autonomous function (18.64s (\pm 13.25)) than just before the TOR (10.70s (\pm 11.6)). No difference was measured between duration conditions in each TOIs.

4.1.2 Sequence B

As for sequence A, the analysis of TOIs (see Fig.2) revealed a difference between the two time windows ($F(1,98) = 4.92$, $p < 0.05$, $\eta^2 = 0.06$ (small)). Indeed, the average duration spent on driving environment in the post_MtoA time window is 14.26 (\pm 11.72) while the average duration is lesser for the period just before the TOR notification (9.04 (\pm 11.83)).

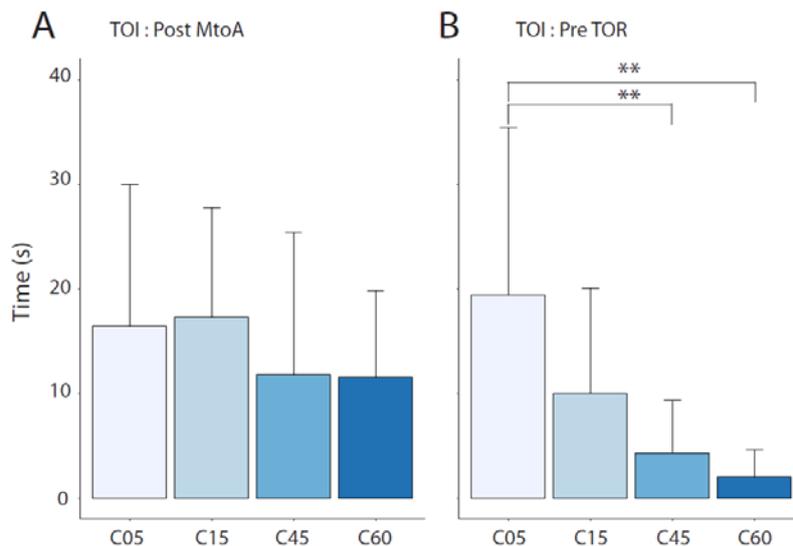


Figure 3. Bar plots illustrate mean total duration of visits (seconds) for the AOI driving environment in Sequence B, for each group and each temporal window (3A: Post_MtoA, and 3B: pre_TOR). Vertical black lines illustrate standard deviation for each mean value. ** for $p < 0.01$.

Statistical analysis also revealed a significant difference between groups for the TOI pre_TOR (see Table 1). Tukey post-hoc analysis indicated that in the TOI pre_TOR, a longer total duration of visits was observed in the shorter duration group (C05), than in the two longer duration groups (C45 and C60) (see Fig.3B).

5. Conclusion

Results showed that during the same drive, the more time passes, the more participants relinquished visual control on the driving environment, as shown by the differences for the two times of interest. Further, we showed a specific effect of autonomous driving time conditions in the second exposure, where subjects who were exposed to short durations remained more attentive to the driving environment than participants who were exposed to long durations. These differences in driving environment monitoring underlined a change in the strategy of the participants, potentially reflecting an increased trust in the system. We suggest that participants try to anticipate the occurrence of the TOR in sequence B, acting as if the TOR would appear within the same timing as in sequence A (albeit the uncertainty providing by the experimentalist, no information of the duration condition). Nevertheless, these modulations, especially for long durations groups could also be due to a potential effect of the secondary task which could increase the workload and then redirect attentional load of participant as time goes by. Further analysis of ocular activity on the area of interest relative to the non-driving secondary task could provide more information to this interpretation.

6. Acknowledgement

This study is a part of the OpenLab agreement “Automotive Motion Lab” between Groupe PSA and Aix-Marseille University and issued from the “Back into the loop Project” supported by the Fondation MAIF pour la Recherche.

References

- Association, W. M., & others. (2001). World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *Bulletin of the World Health Organization*, 79, 373.
- Bahner, J. E., Hüper, A. D., & Manzey, D. (2008). Misuse of automated decision aids : Complacency, automation bias and the impact of training experience. *International Journal of Human Computer Studies*, 66(9).
- Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation research part F: traffic psychology and behaviour*, 18, 47-57.
- Carroll, J. M., & Olson, J. R. (1988). Mental models in human-computer interaction. *Handbook of human-computer interaction*, 45-65.
- Du, N., Yang, X. J., & Zhou, F. (2020). Psychophysiological responses to takeover requests in conditionally automated driving. *Accident Analysis & Prevention*, 148, 105804.
- Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human factors*, 37(1), 65-84.
- Gold, C., Körber, M., Lechner, D., & Bengler, K. (2016). Taking over control from highly automated vehicles in complex traffic situations: the role of traffic density. *Human factors*, 58, 642–652.
- Feldhütter, A., Gold, C., Schneider, S., & Bengler, K. (2017). How the duration of automated driving influences take-over performance and gaze behavior. In *Advances in ergonomic design of systems, products and processes* (pp. 309-318). Springer, Berlin, Heidelberg.
- Kaber, D. B., & Endsley, M. R. (1997). Out - of - the - loop performance problems and the use of intermediate levels of automation for improved control system functioning and safety. *Process safety progress*, 16(3), 126-131.
- Lounis, C., Peysakhovich, V., & Causse, M. (2021). Visual scanning strategies in the cockpit are modulated by pilots’ expertise: A flight simulator study. *PLoS one*, 16(2), e0247061.
- Merat, N., Seppelt, B., Louw, T., Engström, J., Lee, J. D., Johansson, E., ... & Keinath, A. (2019). The “Out-of-the-Loop” concept in automated driving: proposed definition, measures and implications. *Cognition, Technology & Work*, 21(1), 87-98.
- Metz, B., Wörle, J., Hanig, M., Schmitt, M., Lutz, A., & Neukum, A. (2021). Repeated usage of a motorway automated driving function: Automation level and behavioural adaption. *Transportation research part F: traffic psychology and behaviour*, 81, 82-100.
- Naujoks, F., Purucker, C., Wiedemann, K., & Marberger, C. (2019). Noncritical State Transitions During Conditionally Automated Driving on German Freeways: Effects of Non-Driving Related Tasks on Takeover Time and Takeover Quality. *Human factors*, 61, 596–613.
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, 30(3), 286-297.
- Yoon, S. H., Kim, Y. W., & Ji, Y. G. (2019). The effects of takeover request modalities on highly automated car control transitions. *Accident Analysis & Prevention*, 123, 150–158.

Age effects on the Box Task combined with a Detection Response Task

Tina Morgenstern^{1*}, Daniel Trommler¹, Ines Karl², Frederik Naujoks², Josef F. Krems¹, Andreas Keinath²

¹ Cognitive & Engineering Psychology, Chemnitz University of Technology, Chemnitz, Germany

² BMW Group, Munich, Germany

*tina.morgenstern@psychologie.tu-chemnitz.de

There are several standardized test methods to assess the distraction potential of secondary task engagement while driving. One relatively new method is the Box Task combined with a Detection Response Task (BT + DRT). This method has the potential to distinguish between different dimensions of driver distraction. While the DRT is being implemented as an ISO standard, the BT is not yet standardized. There are open questions that have to be clarified in preparation for a standardization. One important issue is potential unwanted variation in BT performance due to the age distribution within a sample. Therefore, the present analysis investigates if age has an effect on BT performance. Fifty-two participants completed an easy and difficult version of a cognitive as well as visual-manual secondary task while simultaneously performing the BT + DRT. Age differences could be shown for BT performance: Compared to younger participants, older participants produced significantly higher variabilities in box position and box size across all secondary task conditions. The results should be considered in future studies.

1. Introduction

The Box Task combined with a Detection Response Task (BT + DRT, see Hsieh & Seaman, n. d.) is an easy-to-use test method to assess potential distraction effects caused by secondary task engagement while driving (e.g., using in-vehicle infotainment systems). The method is based on the Dimensional Model of Driver Demand (Young, Seaman, & Hsieh, 2016). Thus, lateral and longitudinal vehicle control is related to visual-manual demand, while event detection is associated with cognitive demand. In the BT + DRT method, the BT is intended to capture visual-manual and the DRT cognitive demand. This allows a distinguishability in terms of different distraction dimensions (Morgenstern et al., 2020a, Morgenstern et al., 2020b).

While the DRT has been standardized by the ISO (ISO 17477, 2016), a standardization of the BT in combination with the DRT is still pending. There are open questions that need to be addressed to ensure comparability of results across studies. For example, sample selection, such as the age distribution, might have an effect on performance parameters, and thus, leading to unwanted biases.

Previous research indicated age effects on driving performance during secondary task engagement. For example, Merat, Anttila, and Luoma (2005) found that older drivers tend to have a closer car following, higher speed variations and lower lane keeping performance while performing secondary tasks compared to drivers of average age. Bunce, Young, Blane, and Khugpath (2012) reported in their study that older drivers show a higher variance in headway and lateral lane position, resulting in greater inconsistencies in driving performance, than younger drivers. Moreover, they found that the driving tasks are mentally more demanding for older drivers compared to younger drivers.

The objective of the present paper was to examine if there are differences in BT performance between younger and older participants. This might have implications for sample characteristic requirements.

2. Method

2.1 Participants

Overall, 52 participants (26 females, 26 males) with a mean age of 44 years ($SD = 20.19$) participated in the study. Twenty-five participants were younger than or equal to 35 years, 26 participants were older than or equal to 55 years. One participant with an age of 39 years was excluded for age analyses. Hence, 51 participants were included in the analyses.

2.2 Material

2.2.1 BT + DRT

For the present study, we used a PC-based version of the BT. The BT is a continuous tracking task in which participants have to keep a dynamic box within two boundaries using a steering wheel for lateral and a gas pedal for longitudinal box control (see Trommler et al., 2021). In parallel, participants need to respond to vibration stimuli presented in random intervals by pressing a button on the steering wheel (see ISO 17488, 2016).

2.2.2 Secondary tasks

Two secondary tasks were used in an easy and difficult version – a *counting task* (see e.g., Petzoldt & Krems, 2014) and the *Surrogate Reference Task* (SuRT; Mattes & Hallén, 2009). During the counting task, participants had to count forwards in steps of two (easy version) versus backwards in steps of seven (difficult version) starting from a specified number (e.g., 212). During the SuRT, participants had to identify a larger white circle (target) within a number of smaller white circles (distractors). The easy and difficult versions of the SuRT differed in the size of targets and distractors.

2.3 Procedure

The experimental session consisted of a practice trial (i.e., practicing the BT + DRT), a baseline condition (i.e., performing the BT + DRT without secondary task engagement) and four dual-task conditions (i.e., performing the BT + DRT with secondary task engagement). The dual-task conditions were balanced. Participants were instructed to perform the BT + DRT as safely as possible while simultaneously engaging in the secondary tasks. Each trial lasted three minutes.

3. Results

Mixed ANOVAs regarding the mean standard deviation of box position (SDLatP) and box size (SDLongP) were conducted. The between-subjects factor was age group (i.e., ≤ 35 years ($N = 25$) and ≥ 55 years ($N = 25$)); one participant had to be excluded after visual analysis of the boxplots ($>$ three interquartile ranges regarding the number of lateral and longitudinal errors during the baseline condition). The within-subjects factor was secondary task condition. Mauchly's tests indicated violations for the assumption of sphericity for the factor secondary task. Therefore, Greenhouse-Geisser corrected degrees of freedom are reported.

3.1 Mean standard deviation of box position

Generally, both younger and older participants showed the highest variability of box position during the difficult SuRT condition, the lowest during the baseline condition (see Fig. 1). Box position variability differed significantly across secondary task conditions in both age groups ($F_{\leq 35\text{years}}(2.215, 53.159) = 20.110, p < .001, \eta_p^2 = .456$; $F_{\geq 55\text{years}}(2.641, 63.372) = 13.824, p < .001, \eta_p^2 = .365$). The mixed ANOVA revealed no significant interaction between age group and secondary task conditions ($F(2.730, 131,058) = 1.169, p = .322, \eta_p^2 = .024$), indicating that participants of the age group ≤ 35 years produced less variability in box position across all secondary task conditions compared to participants of the age group ≥ 55 years. There were significant differences between younger and older participants across all secondary task conditions (see Fig. 1).

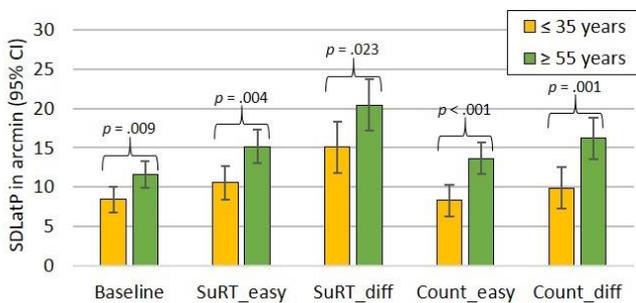


Fig. 1. Mean standard deviation of box position (SDLatP) across secondary task conditions depending on age group.

3.2 Mean standard deviation of box size

Similar results were found for box size variability. Both younger and older participants produced the highest variability of box size during the difficult SuRT condition, the lowest during the baseline condition (see Fig. 2). There were significant differences in box size variability between

secondary task conditions in both age groups ($F_{\leq 35\text{years}}(3.741, 89.795) = 12.723, p < .001, \eta_p^2 = .346$; $F_{\geq 55\text{years}}(2.833, 67.980) = 17.146, p < .001, \eta_p^2 = .417$). The mixed ANOVA revealed no significant interaction between age group and secondary task condition ($F(3.173, 167.730) = 1.814, p = .137, \eta_p^2 = .036$). Hence, less box size variabilities were observed for younger compared to older participants. Except for the baseline condition, there were significant differences between the age groups across the secondary task conditions (see Fig. 2).

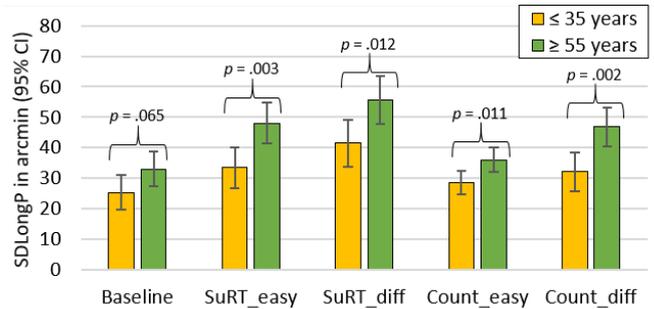


Fig. 2. Mean standard deviation of box size (SDLongP) across secondary task conditions depending on age group.

4. Conclusion

The analyses revealed a clear age effect on the BT performance. This is consistent with research findings of previous studies investigating differences in driving performance depending on drivers' age (Bunce et al., 2012; Merat et al., 2005). For example, in the study of Merat et al. (2005), older drivers showed a lower lane keeping performance and higher speed variations during dual-task trials compared to drivers of average age. In the present study, the older age group was associated with a significantly higher variability in box position (representing lane keeping performance) as well as box size (representing headway to the lead vehicle) compared to the younger age group. This was observed for almost all task conditions. However, there were no interactions between participants' age and secondary task condition, indicating that differences between secondary task conditions within an age group were comparable between the age groups. The results might have implications for sample selection in future studies: When investigating the *absolute* distraction potential of a secondary task (e.g., using a new in-vehicle technology), an age-balanced sample should be used. However, when assessing the *relative* distraction potential (e.g., the distraction potential of a secondary task compared to others), age seems not to play a major role.

5. Acknowledgments

This research was supported by the BMW Group.

References

- Bunce, D., Young, M. S., Blane, A., & Khugpath, P. (2012). Age and inconsistency in driving performance. *Accident Analysis and Prevention*, *49*, 293–299. <https://doi.org/10.1016/j.aap.2012.01.001>
- Hsieh, L., & Seaman, S. (n.d.). Evaluation of the Two-Dimensional Secondary Task Demand Assessment Method.

Unpublished report, Department of Communication Sciences and Disorders, Wayne State University.

ISO 17488 (2016). Road vehicles — Transport information and control systems — Detection-response task (DRT) for assessing attentional effects of cognitive load in driving.

Mattes, S., & Hallén, A. (2009). Surrogate distraction measurement techniques: The lane change test. In M. Regan, J. Lee, & K. Young (Eds.), *Driver Distraction* (pp. 107–122). CRC Press.

Merat, N., Anttila, V., & Luoma, J. (2005). Comparing the driving performance of average and older drivers: The effect of surrogate in-vehicle information systems. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8, 147–166.
<https://doi.org/10.1016/j.trf.2005.04.011>

Morgenstern, T., Trommler, D., Forster, Y., Naujoks, F., Hergeth, S., Krems, J. F., & Keinath, A. (2020a). Measuring Driver Distraction with the Box Task – A Summary of Two Experimental Studies. In H. Krömker (Ed.), *HCI in Mobility, Transport, and Automotive Systems. Driving Behavior, Urban and Smart Mobility* (Vol. 12213, pp. 51–60). Springer International Publishing.
https://doi.org/10.1007/978-3-030-50537-0_5

Morgenstern, T., Wögerbauer, E. M., Naujoks, F., Krems, J. F., & Keinath, A. (2020b). Measuring driver distraction – Evaluation of the box task method as a tool for assessing in-vehicle system demand. *Applied Ergonomics*, 88, 103181. <https://doi.org/10.1016/j.apergo.2020.103181>

Petzoldt, T., & Krems, J. F. (2014). How does a lower predictability of lane changes affect performance in the Lane Change Task? *Applied Ergonomics*, 45(4), 1218–1224. <https://doi.org/10.1016/j.apergo.2014.02.013>

Trommler, D., Morgenstern, T., Wögerbauer, E. M., Naujoks, F., Krems, J. F., & Keinath, A. (2021). The box task - A method for assessing in-vehicle system demand. *MethodsX*, 8, 101261.
<https://doi.org/10.1016/j.mex.2021.101261>

Young, R., Seaman, S., & Hsieh, L. (2016). The Dimensional Model of Driver Demand: Visual-Manual Tasks. *SAE International Journal of Transportation Safety*, 4(1), 33–71. <https://doi.org/10.4271/2016-01-1423>

“Who’s Got the Remote Control?” Understanding Driver Distraction and Inattention in the Context of Teleoperation and the Passenger Experience

Madelaine Coffey*, Madeline Hallewell, David R. Large, Catherine Harvey, Gary Burnett

¹ Human Factors Research Group, University of Nottingham, Nottingham. UK

*maddiecoffey3@gmail.com

Abstract: The remote operation of automated vehicles (‘teleoperation’) has been posited as a potential solution for situations in which human intervention is required, but creates new challenges for ‘driver’ distraction and inattention. Guided by the critical decision method, we conducted an interview study involving 12 experts and practitioners in this nascent field and present a brief overview of emerging areas that require further research attention, in particular regarding workstation design requirements for remote operators to avoid or minimise distraction.

1. Introduction

Fully automated vehicles (AVs) are expected to ameliorate issues of driver distraction and inattention. Indeed, if the vehicle occupant is no longer able to drive (i.e., there are no primary vehicle controls), then, by definition, the vehicle occupant cannot be distracted from the task of driving. However, it is perfectly feasible that AVs may reach the limit of their operational design domain (ODD), be presented with an unexpected, emergency situation outside the scope of the control algorithms, or encounter an unexpected technical malfunction (Kalaiyarasan et al., 2021). In these situations, teleoperation has been proposed as a viable mechanism to provide remote control of the vehicle, for example, to manoeuvre it at low speed to a safe location. In effect, this reassigns the mantle of ‘driver’ to a remote operator (RO), who has hitherto not been actively involved in the control of the vehicle or even present therein. Understanding the issues this presents to the RO, and how to deal with potential distraction and inattention that may ensue, is therefore important.

2. Background

2.1 Future, Mobility-as-a-Service (MaaS) AVs

Future MaaS-AVs are expected to operate as part of an integrated transport system, whereby passengers will be able to book or hail a vehicle, but there will be no driver present (so-called, “robotaxis”). Kalaiyarasan et al. (2021) predict that even for a fully operable service, unforeseen edge-cases and emergency situations may still occur that fall outside the normal mode of operation. In such situations, teleoperation has been proposed as a means to *remotely control* these vehicles; this could extend from limited path guidance (e.g., manoeuvring around roadworks) to full remote driving at low speed or even high speed (Economic Commission for Europe (ECE), 2020). In addition, the ECE (2020) identify two further potential categories of remote support and control, namely, *remote assistance*, whereby the service provider offers support and breakdown assistance, and *remote management*, analogous to air traffic control, in which the remote operator assists when the vehicle requires authority to move or deviate from a prescribed path.

Notwithstanding the technical challenges these present, *remote control*, in particular, also creates a new form

of driver distraction and inattention – notably for a driver who is not even located within the vehicle, thus creating a new research agenda. Further consideration must also be given to the safety and wellbeing of any incumbent passengers, who may inadvertently create additional situations requiring remote intervention, such as the need to travel beyond the ODD in a medical emergency. With these factors in mind, the current interview study sought to uncover the needs of ROs who will be tasked with monitoring AVs and may be required to intervene in vehicle control, management or guidance. In particular, we sought to explore issues pertaining to RO workstation design and its impact on distraction and inattention.

2.2 Remote Operation of AVs

In their vision of robotaxi teleoperation, Zhang (2020) suggests that a teleoperation workstation might mimic the in-car driver experience, enabling the RO to “see, hear and feel” the car and its surroundings. With the additional sensors present on AVs (e.g., LIDAR etc.), combined with artificial intelligence (AI) systems, ROs could subsequently be presented with much more information than the standard driver. Delivering such information succinctly to a human operator is likely to be difficult: cognitive overload may subsequently ensue (Mutzenich, Durant, Helman, & Dalton, 2021). Mutzenich et al. (2021) also argue that the teleoperation of AVs presents challenges relating to out-of-the-loop (OOTL) syndrome, latency, embodiment and workload. They note that ROs will encounter a potentially hazardous delay whilst attaining adequate situational awareness when they have been given a vehicle to control.

There are already some general user-requirements outlined in the literature regarding teleoperation (Georg & Diermeyer, 2019; Graf & Hussmann, 2020; Kettwich, Schrank, & Oehl, 2021). However, at the time of writing, there is limited knowledge surrounding the AV-robotaxi teleoperator role specifically, in which the operator must consider the needs of the driving task as well as that of the human passenger in the vehicle, who may be showing varying levels of unease depending on the circumstances, or may require other assistance or support.

3. Method

Guided by the Critical Decision Method (Klein et al., 1989), we conducted an interview study involving 12 experts.

Participants either held direct responsibility for the teleoperation of AVs or similar vehicles/equipment, were working in companies that are developing teleoperation technologies/services, and/or working on relevant research projects. Participants comprised existing contacts of the ServCity team (<https://www.servcity.co.uk/>) and Centre for Connected and Autonomous Vehicles (CCAV). Interviews aimed to uncover potential issues associated with the remote operation of robotaxis (with a focus on workstation design and vehicle control), whilst maintaining the wellbeing and a positive user experience for the vehicle occupant.

4. Results

Interviewees highlighted many new issues associated with the role of RO in the context of MaaS-AVs, and workstation design, which are thematically grouped and summarised below. The findings represent the experiences and opinions of the 12 experts (identified as P1 to P12, to whom specific comments are attributed) and are intended to provide an outline and inspiration for future research studies.

4.1 Accessibility

Accessibility of remote operation has not received much research attention, although there are international standards for workplace accessibility. The potential for multimodal feedback of increased sources of data (e.g., LIDAR sensor data) could potentially open up the RO role to people previously excluded from driving tasks (P5, P6).

4.2 Working patterns and demands

Human capabilities and limitations must be considered within the creation of the RO role, in particular the ability to maintain vigilance over long periods. Shifts should be carefully arranged to protect employee health whilst making the most of their abilities (P4, P5, P7).

4.3 Human Machine Interfaces

Visual display screens are expected as a minimum (P1, P7). There was also some suggestion that augmented and virtual reality equipment might be beneficial in certain situations (P5), in which case, the visual quality should be as high as possible. Icons and auditory alerts could replace text for some operations (P4).

4.4 Physical setup of workstations

International standards already mandate some aspects of workstation design (e.g., desk height). However, there is no consensus on the number of screens and whether the RO needs access to a full 360° view (P1, P9). It may be necessary to have multiple communication channels available (P1, P2, P3). ROs might need to be able to see and collaborate with colleagues/supervisors in the room with them (P3).

4.5 Handover

There must be a clear, unambiguous process of handover (P1, P9). The system could check that the RO is ready using driver monitoring systems. There were mixed views about whether passengers always needed to be informed that the vehicle is being operated remotely, but it was felt that it would be useful for passengers to be aware in certain situations, such as in an emergency (P1, P9).

4.6 Latency

Network latency is already recognised as a potential problem, and the network infrastructure must be designed to ensure that this is kept to a minimum. It was suggested that a latency of 100 milliseconds or more can impair performance at higher road speeds (P5). Minimal risk manoeuvres could be utilised where there is a risk of high latency (P1, P9, P10).

4.7 Role of the teleoperator

The exact job description of a RO is still a topic for discussion, with some suggesting it may comprise a low-skilled role simply monitoring vehicles and selecting intervention options from a pre-determined list (P7), whereas others suggested that it should be a high-skilled role in which the operator can take over control of the vehicle even whilst it is moving (P1, P4 P12).

4.8 Workload

Workload for a RO is likely to be variable, with long periods spent monitoring and shorter periods of high cognitive/physical activity. Moreover, ROs may be responsible for supervising multiple vehicles (P10) or acting as a guide vehicle for a convoy (P12). They will therefore need to be able to cope with demanding and stressful situations as well as overcoming boredom (P1). The road situation (e.g., other vehicles/pedestrians, traffic signals) will also contribute the level of workload experienced (P8).

4.9 Situation awareness

ROs will need support to establish situation awareness, potentially requiring guidance from the automated system (P9). Providing information before handover, along with suggested courses of action might alleviate the impacts of reduced or absent situation awareness (P7, P10).

4.10 Information needs (including training)

New and specific training will be needed (P01, P09, P10). ROs may not be required to hold an existing driving licence, particularly given the different roles they may fulfil (remote control, assistance and/or management) (P10, P12), but those who are in control of a vehicle on a road must have a licence to drive in the country in which they are “driving” (P09, P12). Visual information will be most important, but can be supported by audio, LIDAR sensor data, haptic feedback and route-related information (P1, P2).

4.11 Supporting passengers

ROs will need audio communications with passengers, but it is unclear whether ROs or customer service agents will be responsible. Passengers will likely need support and reassurance from the RO (comments from all participants).

4.12 Other communication

The RO is likely to need a means to contact emergency service operators as well as their colleagues/supervisors, and other road users in the vicinity of the ego vehicle (e.g., other drivers/vehicles involved in an incident) (P6, P8, P10, P12).

4.13 Distraction and inattention of RO

There is arguably a higher risk of distraction for ROs than conventional in-vehicle drivers, as the workstation will enable and sometimes demand the use of multiple interfaces when in charge of a vehicle (or multiple vehicles), for example, to access a team chat to escalate issues or contact emergency services (P9, P10, P11). Moreover, there is the possibility of ROs becoming bored or distracted by mobile phones or other tasks whilst waiting for takeover requests, or whilst supervising a vehicle which, in theory, should be able to drive itself (P06, P10). Driver monitoring systems could be used to assess whether the RO is attending to the situation, or if they are ready to respond to a new task (P04).

5. Discussion and Conclusion

The interviews revealed several key areas relating to the teleoperation of MaaS-AVs, and associated workstation design, that require further research attention. Although interviews were framed within the context of a future, robotaxi service, many of the findings are also applicable to other situations in which the teleoperation of an AV presents as a viable solution.

While the experts interviewed during the study were naturally strong advocates of teleoperation, the concept appears to divide opinion amongst the wider population. Indeed, one could reasonably opine that if an AV was to encounter an unknown situation within its ODD (that subsequently required remote, human intervention), then it is not truly automated (i.e., may require further technical development) and/or the ODD should be redefined. Moreover, the interviews revealed numerous, complex technical issues (network latency, aggregation and visualisation of sensor data, control actuator design etc.) as well as multiple human factors (distraction, workload, situation awareness etc.) that must all be satisfactorily resolved. Consequently, teleoperation may not be the optimal solution in all situations.

Regardless, it remains a popular option in the anticipated, near-future transportation landscape, not least because it encompasses a range of paradigms (remote control, remote assistance, remote management) (ECE, 2020). As such, teleoperation—in the sense of *remote control* of the vehicle—may be most appropriately applied to support the successful implementation and real-world integration of AVs (or applied as ‘last resort’ when all else fails). As the need for remote control of an AV reduces (i.e. the on-board AI systems become more attuned to ‘edge-cases’), teleoperation could provide important, ongoing benefits, such as support and breakdown assistance and/or as a mechanism to provide the necessary authorisation to move or redirect the vehicle from its current path. Within service-led applications, such as robotaxis, ROs could act as the point-of-contact and even as a substitute for an on-board authority figure (see: Dolins, Strömberg, Wong, and Karlsson, 2021), particularly relevant to situations of shared occupancy. The issues highlighted herein are relevant to all these scenarios.

Future human factors research in this area should be directed towards the features of the remote workstation that will help ROs avoid and/or overcome potential distractions, manage their workload and maintain their attention on the task at hand, for example, by providing a sense of realism and embodiment to the (remote) driving task. In ongoing work, these challenges have been aligned with a goal-directed task

analysis, and emerging teleoperator ‘workstation guidelines’ are forthcoming.

6. Acknowledgements

The study was conducted as part of the ServCity project (<https://www.servcity.co.uk/>), which is funded by the UK Innovation Agency (Innovate UK) and Centre for Connected & Autonomous Vehicles (CCAV) (Grant number: 105091).

References

- Dolins, S., Strömberg, H., Wong, Y. Z., & Karlsson, M. (2021). Sharing Anxiety Is in the Driver’s Seat: Analyzing User Acceptance of Dynamic Ridepooling and Its Implications for Shared Autonomous Mobility. *Sustainability*, 13(14), 7828. Retrieved from <https://www.mdpi.com/2071-1050/13/14/7828>
- Economic Commission for Europe (ECE). (2020). Human Factors Challenges of Remote Support and Control A Position Paper from HF-IRADS. Retrieved from Geneva: <https://unece.org/fileadmin/DAM/trans/doc/2020/wp1/ECE-TRANS-WP1-SEPT-2020-Informal-8e..pdf>
- Georg, J.-M., & Diermeyer, F. (2019, 6-9 Oct. 2019). An Adaptable and Immersive Real Time Interface for Resolving System Limitations of Automated Vehicles with Teleoperation. Paper presented at the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC).
- Graf, G., & Hussmann, H. (2020). User Requirements for Remote Teleoperation-based Interfaces. Paper presented at the 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Virtual Event, DC, USA. <https://doi-org.nottingham.idm.oclc.org/10.1145/3409251.3411730>
- Kalaiyaran, A., Simpson, B., Jenkins, D., Mazzeo, F., Ye, H., Obazele, I., . . . Wilford, R. (2021). Remote operation of Connected and Automated Vehicles: Project Endeavour - WP15B. Retrieved from UK:
- Kettwich, C., Schrank, A., & Oehl, M. (2021). Teleoperation of Highly Automated Vehicles in Public Transport: User-Centered Design of a Human-Machine Interface for Remote-Operation and Its Expert Usability Evaluation. *Multimodal Technologies and Interaction*, 5(5), 26. Retrieved from <https://www.mdpi.com/2414-4088/5/5/26>
- Mutzenich, C., Durant, S., Helman, S., & Dalton, P. (2021). Updating our understanding of situation awareness in relation to remote operators of autonomous vehicles. *Cognitive Research: Principles and Implications*, 6(1), 9. doi:10.1186/s41235-021-00271-8
- Zhang, T. (2020). Toward Automated Vehicle Teleoperation: Vision, Opportunities, and Challenges. *IEEE Internet of Things Journal*, 7(12), 11347-11354. doi:10.1109/JIOT.2020.3028766

Different types of human failure causing crashes

Katerina Bucsuhaży, Robert Zuvala

Transport Research Centre, Brno, Czech Republic
E-mail: katerina.bucsuhaży@cdv.cz

1. Introduction

The factors leading to the crash has been widely studied. Inattention has been described as a major problem in road safety and one of the most often factors which contribute to the crash [5], [6]. The presented paper is based on the previous work of Van Elslande [1], [2], who defined the main categories of human failure leading to crashes with a sequential chain of human failure. Regan [3] described that the first defined category – failure in information acquisition – has been perceptual failures (not attentional). In the presented paper, the taxonomy developed by Van Elslande has been adapted and applied to the Czech in-depth database..

2. Data and methods

2.1 Czech in-depth study

Data collection was performed within the research project CzIDAS: Czech In-depth Accident Study [4]. Only crashes with injuries in a defined region of South Moravia in the Czech Republic are investigated. The investigation team realized the analysis immediately after the crash occurred at the crash scene. In-depth investigation includes participant interviews, detailed vehicle and infrastructure documentation and subsequent analysis of the crash. The psychologist focused in the interview on all relevant information related to the causes of the crash and the actual mental and physical condition of the participant.

The presented analysis focused on the information about human failure leading to the crash. The cases were retrospectively qualitatively analyzed. Two researchers independently reviewed cases to determine the type of human failure in each of the crashes. In all used cases were possible to determine human failure at least at one of the crash participants, in some two-participants cases both crash participants failed. The dataset used for analysis contains determined human functional failure of about 1447 personal vehicle drivers, 119 truck drivers, 201 motorcyclists, 88 cyclists and 218 pedestrians..

2.2 Crash causation

The human factor includes all the aspects involved in any activity, either positively or negatively. The human factor failure has been the most common factor which contributes to the crash [7]. To understand the causes of the crash in terms of the human factor, the Van Elslande classification model has been used to describe human failure at various levels. The model described human failure as a sequential, epigenetically evolving process. The classification was used to find and identify the level at which the human decision leading to the crash was made and to better understand the nature of the driver's erroneous actions and how to prevent them. The

global stages of the human malfunction chain potentially involved in crashes are:

- Failure in information acquisition
- Failure in the diagnosis of the situation
- Failure in predicting the situation
 - Failure when deciding to undertake the specific manoeuvre
- Psychomotor failure when performing the action
- Overall failure.

Each of the defined main categories has been subdivided into detailed stages.

3. Results

As evidenced by the results of Czech In-depth Crash Study (CzIDAS) from the number of analysed traffic crashes, the most frequent human failure has been a failure in information acquisition (40 %), where the participants fail in the detection of a potentially critical situation. The road user's attention is improperly aimed, cursory or absent. In the whole crash database can be further identified: the failure in diagnosis (27 %), failure in prediction (9 %), failure in the decision (7 %), psychomotor failure (5 %) and the overall failure (12 %). The frequency representation of human failure has been in ascending order as the information should be processed by road users on a cognitive and somatic level.

As evidenced by statistical testing, the type of failure varied depending on the type of road user. The drivers of personal vehicles and trucks as well as pedestrian more are more likely to have detection failure, the motorcycle drivers or cyclists more frequently fail at diagnosis of the situation. The problems with the diagnosis are based mostly on problems of evaluating physical parameters. The drivers have often problems evaluating infrastructure related difficulties (e.g. curves characteristics and reduced adhesion).

For pedestrians, cursory or hurried information acquisition has been most typical. Drivers of personal vehicles mostly fail due to the information acquisition focused on a partial component of the situation. The detection problems refer to the information acquisition strategy. Also the truck drivers fail most frequently at this level. In comparison to the other drivers, for the truck drivers, the non-detection in visibility constraints conditions has been more frequent. This could be influenced also by vehicle characteristics which could make difficult to detect an important element in a potentially critical situation. The differences in human failure were identified also depending on the age of the road user, the crash type or driver experience..

4. Conclusions

Data from In-depth Crash Analysis provide a comprehensive view of all the factors related to a particular crash and serve to identify the characteristics leading to the crash occurrence. The primary tasks of in-depth study have been to identify how and why the crash occurred and how to prevent a similar situation. The unique findings provided by the in-depth crash analysis are utilized for the definition of priorities and methodological management in the field of road safety, standards and legislative framework modifications and development. Identifying factors leading to human failure can contribute to the targeted direction of preventive activities in the field of road safety and road safety campaigns.

5. Acknowledgments

This article was produced with the financial support of the Ministry of Transport within the programme of long-term conceptual development of research institutions on the research infrastructure.

References

- [1] VAN ELSLANDE, Pierre, et al. Analyzing 'human functional failures' in road crashes. Final report. Deliverable D, 2007, 5.
- [2] VAN ELSLANDE, Pierre; NAING, Claire L.; ENGEL, Ralf. Analyzing human factors in road crashes: TRACE WP5 Summary Report. 2008.
- [3] REGAN, Michael A.; HALLETT, Charlene; GORDON, Craig P. Driver distraction and driver inattention: Definition, relationship and taxonomy. *Crash Analysis & Prevention*, 2011, 43.5: 1771-1781.
- [4] BUCSUHAZY, K. et al. Czech in-depth study. 2021. Transport Research Centre. Brno. Czech Republic. Ministry of Transport
- [5] WUNDERSITZ, Lisa. Driver distraction and inattention in fatal and injury crashes: Findings from in-depth road crash data. *Traffic injury prevention*, 2019, 20.7: 696-701.
- [6] BEANLAND, Vanessa, et al. Driver inattention and driver distraction in serious casualty crashes: Data from the Australian National Crash In-depth Study. *Crash Analysis & Prevention*, 2013, 54: 99-107.
- [7] ALLAHYARI, Teimour, et al. Cognitive failures, driving errors and driving accidents. *International Journal of Occupational Safety and Ergonomics*, 2008, 14.2: 149-158.



DDI 2022 GOTHENBURG

THE 8TH INTERNATIONAL CONFERENCE ON
DRIVER DISTRACTION AND INATTENTION