THE 9TH INTERNATIONAL CONFERENCE ON DRIVER DISTRACTION AND INATTENTION

# DDI 2024 MICHIGAN

University of Michigan, Ann Arbor
October 22–24, 2024

# ABSTRACT BOOK

# INDEX

## Panel presentation



## Session 01 - Countermeasures



## Session 02 - Driver characteristics and individual differences



## Session 03 - Automation 1



## Session 04 - Methods and models



## Session 05 - Driver engagement with technologies



## Session 06 - Driver behavior

# Supporting Driver Attention Holistically: Perspectives from the AHEAD Consortium

Bryan Reimer[1*], Linda Angell[2], Bruce Mehler[1], Lee Skrypchuk[3], Steven Feit[4], Gregory M. Fitch[5], Alexandria M. Noble[6]

[1] Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA, USA
[2] Touchstone Evaluations, Inc., 81 Kercheval Avenue, Suite 200, Grosse Pointe Farms, MI 48236, USA
[3] JLR, Abbey Road, Whitley, Coventry, CV3 4LF, United Kingdom
[4] Honda Development & Manufacturing of America LLC, 21001 State Route 739, Raymond, OH, USA
[5] Google, 1600 Amphitheatre Pkwy, Mountain View, CA USA, USA
[6] CARIAD, Inc., 450 National Ave. Mountain View, CA 94043, USA
[*]reimer@mit.edu

**Abstract: Since the current driver distraction guidelines were developed, the scientific understanding of glance behavior, attention threading, situation awareness, the role of driving context, and other related topics has advanced, based to a significant extent on naturalistic driving research. In addition, vehicle systems have progressed with new forms of external and internal sensing, increased computational capabilities, better screens, greater integration of multi-modal interfaces, driver monitoring, and driver feedback systems. A panel discussion will summarize relevant research and a new conceptual approach for addressing attention management through system design and driver support being developed by the Advanced Human Factors Evaluator for Automotive Demand (AHEAD) consortium. AHEAD is an MIT based, industry-academic pre-competitive collaborative entity, working to build on previous work, while developing an updated approach to driver vehicle interface design, validation, and testing that improves system usability while enabling a foundation for real-time driver attention support. The premise is to build upon existing work, introduce attention centric design, and in real-time assess whether drivers are paying sufficient attention for the current situation. The aim is to leverage technology to promote the rebuilding of situationally relevant knowledge and readiness to respond. This paper summarizes the foundations for the framework and select operational considerations.**

## 1. Introduction

Current global driver-focus and driver distraction guidelines were developed based upon a rich understanding of drivers' interaction with traditional static and largely visual-manual driver vehicle interfaces (DVIs). At the time, the automobile industry and regulators were concerned with the expansion of tasks that could be undertaken while driving and their associated demand on the driver. System manufacturers were just beginning to explore multimodal interfaces and design approaches aimed at mitigating sustained demand. Portable devices (e.g., smartphones) and modern social media were largely yet to influence the connected experience.

While extensive collaborative research had been done (e.g., Angell et al., 2006), limited insight existed on the role of operational context on demand, the benefits and limitations of voice enabled and touchscreen interfaces, the importance of on-road glances, the capabilities of external perception, and the viability of in-cabin sensing to support driver readiness. Modern DVIs are largely multi-modal and disconnected (beyond navigation) from an awareness of the operating context, driver state, and an ability to adapt moment by moment to user needs. Systems have now been deployed with attentional cues designed to draw the eyes to the road (e.g., GM Super Cruise) by leveraging the human's instinctual attraction to motion effects in the periphery.

The Advanced Human Factors Evaluator for Automotive Demand (AHEAD) consortium is an MIT led global industry-academic effort presently consisting of Google, Honda, VW Group, JLR, and Touchstone Evaluations, working as a pre-competitive entity to develop a new conceptual approach for addressing attention management based on historical foundations and new science.

AHEAD's efforts consider the realities of portable electronic use, limitations of current guidelines, a vision towards better DVI design, updated assessment approaches, and a need for safer roads. A panel presentation and this paper builds on earlier work (e.g., Coughlin et al., 2011, Reimer, et al., 2016; Reimer et al., 2022), to share additional details on our vision for supporting driver attention holistically. The framework promotes attention- centric user interface design, the implementation of real-time approaches to rebuilding attention when required, and the use of countermeasures when driver behavior falls outside of acceptable tolerances.

## 2. Historical Development of Past Guidelines

For more than two decades, driver workload and driver distraction have been a subject of concern in the traffic safety arena. This concern began to deepen when electronic devices began to transform the tasks that drivers could undertake while driving. In 2000, NHTSA held an internet forum on the topic of driver distraction. This event triggered many different entities in the U.S. and abroad to act. A number of research programs were initiated with the goal of understanding driver distraction and various related issues. Formative policy discussions began taking place.

Much of the effort was on a "device-oriented" perspective and focused on assessing *the level and types of workload demands* that new vehicle subsystems placed on drivers – based on the notion that the demands on the driver should not be excessive.

While it was understood at that time that fully addressing distraction entailed issues of attention ***in addition to*** issues of workload – in the early 2000's the technology for attentional cuing and related attention-related countermeasures was far from ready for deployment. Therefore, to attempt to limit driver distraction, initial steps were made with a focus on driver workload. There was hope that at some future point it would become possible to also begin addressing issues of driver attention.

The type of approach taken in the early frameworks offered a number of practical advantages to manufacturers from the point of view of ***developing*** DVIs and their subsystems (by placing an emphasis on the *design of tasks and interface elements* that a manufacturer could influence and optimize through design and engineering). However, these approaches were also **highly constrained**. Specifically:

1. A secondary task was treated as **a single unit of analysis**, as an epoch of time **removed** from the larger continuum of driving – and **removed** from a consideration of varying concurrent demands of the driving task.

2. ***Fixed limits*** were placed on the "amount" of demand that a secondary task could impose on a driver during the slice-of-time that a single task took – and these limits were placed on each *individual type of demand* considered *separately* from all others (e.g., visual demand considered separately from auditory or cognitive demand).

3. The fixed **limits were *invariant* across driving scenarios and conditions** that themselves would typically vary in the amount of attention they required from drivers. Instead of testing across varying, representative driving scenarios, test methods assessed whether or not tasks interfered with driving in a *single standardized type of driving scenario* (which was a car-following scenario in a low-demand driving environment on a straight road). This was intended to reflect the type of setting that at the time had most often been associated with the conditions under which distraction-related crashes had been observed, based on a synthesis of the research reported by Bents (2000), Hendricks, Fell, & Freedman (2001), Stutts et al. (2001), and Wang, Knipling, and Goodman (1996).

4. These limits were rendered on a ***dimension-by-dimension basis*** (e.g., visual demand separately from other dimensions) – **and no means was typically provided for considering conjoint or interleaved demands of "multiple types"** on the driver by a task. Further, *no means for combining* results across tests or across resource dimensions into an overall measure of task load to obtain a holistic "big picture" of a task's effect on the driver and their driving performance was typically provided in the early frameworks. As a result, the early *frameworks had difficulty handling the evaluation of tasks that were complex and placed multiple types of demands on the driver* – e.g., multi-modal demands in rapid succession, or the intricate threading of task elements over time, or even the management of task elements and modalities concurrently. If multimodal tasks were evaluated in this early period, the existing methods required such tasks to be evaluated *multiple times* (each time using a different evaluation methodology for each type of resource demanded – e.g., glance measurement for visual demand, Detection Response Task for cognitive demand, etc.). *There were a few exceptions which provided an overall metric for performance as a function of any task load (whether on a single dimension or multiple dimensions) – for example, the lane change test (Mattes (2003), Mattes & Hallén (2009)), Burnett et al., (2013) and the box test (which is still under study – e.g., Morgenstern et al (2020)). However, these two methods have thus far been ancillary to the more common practice of evaluating tasks dimension-by-dimension.*

5. In addition, the early approaches to distraction focused solely on **preventing excessive demand on drivers** – *rather than* – **on supporting drivers as they try to optimize their level of attentiveness to driving**. These two objectives are very different. Achieving one of them (e.g., *ensuring that secondary task demand is not excessively high*) does not necessarily mean that the other will also be achieved (*i.e., that drivers will be effectively-supported in attending to the road when-and-where they need to be*).

**Thus, the early frameworks typically did *not* consider whether drivers were attending to the road in an adequate manner** (e.g., an adequate amount, at appropriate times, and with adequate levels of attentional arousal). This was largely because, at the time, the tools were not yet available for conducting naturalistic driving studies - and virtually no data were available regarding when drivers chose to initiate tasks (under what conditions of driving) – so questions about how drivers managed attention over time under natural conditions could simply not be examined.

6. Further, in the early frameworks, ***conditions of underload*** **were often not considered at all** (even though during states of underload, monotony, and boredom, drivers often *initiate* secondary tasks as a means of increasing and/or optimizing their levels of attentional arousal). Indeed, conditions of "increasing workload" were almost always assumed to be *undesirable* – and treated as such. Yet, published findings now suggest that when attentional arousal is low, performance can sometimes be improved with a heightening of arousal – and show that increases in task loads of certain types can improve overall performance.

Thus, on the one hand, each of the six constraints itemized above could be seen as a shortcoming of the early distraction frameworks. **However, we see them differently. We see them as opportunities to advance the state of the art** as **many things have changed** since the first guidelines limiting distraction were formulated. Now, through naturalistic driving studies, **much more is understood about driver behavior** "in the wild." In addition, **technology has advanced** along several dimensions – and now offers the capability to adapt the user-interface, and to offer new types of support to drivers in real-time.

## 3. AHEAD's Scientific Contributions in Support of a Broader View of Glance Behavior & Attention

A major aspect of the AHEAD perspective is a shift in focus from the potentially narrow concept of distraction to a broader consideration of how a driver's attention is distributed over time. This approach asks whether the driver has been attending to the driving task, including the surrounding driving context, sufficiently to safely carry out the immediate task and maintain a level of situation awareness to anticipate and respond to changing events and demands as they emerge.

In terms of visual attention, this perspective considers not only 'distracting events' that take the driver's eyes off the road, but also whether the driver's pattern of glances back to the road are of sufficient duration to re-establish and maintain appropriate situation awareness. The concept of an 'attention buffer' was originally introduced by Kircher and Ahlstrom (2009). As adapted and extended by AHEAD, this view of attention argues that a driver's awareness of the details of the driving scene degrades as they look off-road such that not only does the risk of missing a critical event taking place on-road increase the longer one looks off-road, but maintenance of awareness of the details of the road scene degrade as well, leaving the driver less prepared to respond to events when they return their gaze to the road. Equally critical, this model of attention argues that it takes time once the driver looks back to the road to re-establish a comprehensive picture of the driving environment. Brief 'check' glances back to the road may or may not be of sufficient duration to detect a "bottom-up" stimulus such as brake lights coming on in a lead vehicle. Further, relatively long on-road glances are required in complex driving conditions to reacquire a level of awareness that allows a driver to anticipate emerging conflicts in the details of the road scene ("top-down" processing) and thus act to avoid conflicts before they become safety critical. Consequently, a more comprehensive assessment of attention needs to account for the pattern and duration of on-road glance behavior (and, ideally, the driving context).

AHEAD research has demonstrated that an attention algorithm that considers how a driver threads together both on and off-road glances can differentiate relative safety risks in naturalistic datasets that cannot be differentiated using just off-road glance metrics (e.g., Seaman, et al., 2017; Seppelt, Seaman, Lee, et al., 2017; Seppelt, et al., 2018). AHEAD efforts have explored refinements to the base rules of the initial buffer concept, particularly as regards the reestablishment of situation awareness through on-road glance characteristics as well as other features (e.g., Seppelt, Seaman, et al, 2017; Seaman, et al., 2021). These findings argue for both respecting prior work on the safety significance of off-road glance behavior and the importance of DVI design that considers on-road glance behavior in support of driver situation awareness.

In addition to the published work referenced above, AHEAD has explored the potential for further refinements in the study of on-road glance behaviour to detect divided attentional states such as those associated with high cognitive load or mind-wandering. While technical challenges are currently present (e.g., Wang et al., 2014; Ding et al., 2023), practical implementations are not necessarily far off and can easily be incorporated within the AHEAD framework.

## 4. Implications & Motivation for a New Framework Focused on Driver Attention Support

The availability of new technologies not only creates a need for human centred driver attention management methods, they also enable it. Moreover, system design may benefit from taking a functional approach to driver attention that focuses on a holistic (system-wide) view of the net impact of all sources of demand (primary and secondary, under all assistance levels, and the role of operating context) on safety. AHEAD sees this holistic approach focusing on driver attentional support to promote situation awareness across three interrelated concepts (first elaborated by Angell (2012) and adopted by AHEAD):

- Managing task workload within a zone of acceptability.
- Preventing interference with natural attention allocation strategies and preventing disruptions of the driver's attention functions.
- Supporting a driver's focus when capabilities are limited, they're having difficulty, or something unexpected occurs.

Driver Attention Support is about helping drivers supply sufficient attention for the current driving situation. This can be accomplished through a combination of system design to mitigate workload and protect attention, as well as real-time adaptions that are now increasingly feasible to help ensure that the attention a driver supplies meets or exceeds the attention the driving task requires at a given time, so that drivers are well positioned to respond to developing events. As noted, technical developments have increased capabilities to estimate relative required attention using data on:

- Driving task demands - assessed using vehicle and infrastructure sensors (e.g., speed, map data of congestion/design, camera/radar/ lidar, user-generated content, weather data, and traffic signal SPaT).
- ADAS capabilities - accounting for effectiveness in supporting driving, and reducing crash risk, assessed using FOT studies and safety benefit estimates.
- Driver capabilities - information (e.g., a parental control identifying a novice driver or other historical data) indicating an attention challenged driver.

AHEAD proposes that a new model for DVI design and evaluation be considered that promotes an attention-centric approach. In cases where required attention and supplied attention are unknown, a case can be made for defaulting to current distraction guidelines. However, when the required attention and/or supplied attention can be estimated, extensions to the current guidelines are developed to enable adaptable DVIs that, in real-time, support rebuilding attention as required. Furthermore, countermeasures may be used when driver behavior falls outside of acceptable tolerances (e.g., texting using a personal electronic device). Core interrelated topics are discussed in the following sections.

### *4.1 Attention-Centric Driver Vehicle Interface Design*

One of the most important factors associated with task completion in the vehicle is how well the system is designed in the first place. Following a human-centred approach to design and trying to find the most appropriate interfaces within the context of use is fundamental to achieving a simple and satisfying experience for all vehicle users. The focus on how a driver uses their attention in relation to the demand generated within the driving context is the basis for the AHEAD design approach, and hence signifies a change in direction from existing guidelines that tend to shy away from specific recommendations around interface design.

AHEAD recommends that as an industry we focus more on: simplification of content (Rosenholtz, et al., 2011), avoiding attention-based traps (*i.e., visual search*, Scott, 1993), and using appropriate multi-modal methods of interaction (Schnelle-Walka & Radomski, 2019). Design must be based on empirical findings, as often theory or rationale-based assumptions do not translate to the actual real-world user behavior. Therefore, it is critical to test and compare solutions, in context, to understand how they impact attentional behavior.

**Content simplification** starts with understanding the fundamentals of task design at a system level to make sure that achieving important functional goals are simple and short in duration. For example, surfacing frequently used tasks, or allowing users to configure favourite options to appear more prominently. In the display design process, it is valuable to assess how much content it contains and how deep the structures go, and then reassess how much of the functionality is absolutely necessary.

**Avoiding attention-based traps** starts with the understanding that humans will search for visual information that matches their goal (Wolf & Horowitz, 2017). Therefore, content heavy information displays may naturally lead to long periods of visual search and increased glance durations, thus increasing the chances of supplied attention being compromised. The combination of different interfaces and task types lead to tasks that take longer (*such as typing*) or have no clear end point (*scrolling lists, deep menus*) are examples of interface strategies that can lead to attention-based traps (Large, et al., 2019). Equally, there are also examples of design approaches that automatically trigger glances back to the road, such as contextual cueing (Chun, 2000). Another form of an attention-based trap is when messages are presented at inopportune times. A mis-timed glance during a demanding situation could lead to a misbalance of attention. As discussed later in the countermeasures section, building in mechanisms to schedule feedback at appropriate times is an important attention-centric strategy.

Finally, using **appropriate multi-modal methods** of interaction can help users complete challenging tasks. The difference between good or bad performance is usually good or bad design. For example, determining when visual-manual interaction is necessary and needed as opposed to other modalities is very important. The balance between physical and digital control execution is particularly sensitive. The digital space offers huge variety and flexibility but also significantly increases the potential for bad design. Consider how voice can help or be used proactively rather than letting a user decide which way they want to interact. Utilizing AI methods, can the vehicle itself predict the user's goal and proactively offer the most attention centric way to interact?

Making an optimized design decision amongst a catalogue of potential solutions is part of the challenge. A lot of work, research, and empirical evidence is still required to understand how to support driver attention naturally and, hence, industry should focus on consistent, repeatable interaction methods that target a reduction of visual demand in its entirety.

### 4.2 Required and Supplied Attention

The goal is to build upon previous work and leverage new technologies and research in constructing a driver attention framework that helps support driver focus. As such, given the proliferation of information technology and embedded sensors, it is increasingly feasible to estimate the required attention of the driving task that is used a priori to inform design assumptions and/or on a moment-to-moment basis in to inform adaptive systems. Supplied attention by the driver is a function of the visual attention to the road, cognitive attention to the road, and manual control of the vehicle. Supplied attention over time is critical for building and maintaining a model of the driving situation (see Section 3). Supplied attention can be estimated in various ways based on driver behavior, such as using taps on the display, in-vehicle cameras, and perhaps even voice interaction. Similar to required attention, supplied attention could be estimated a priori (for design) or in real-time to dynamically estimate or forecast the driver's state.

AHEAD sees significant potential in the development of adaptive DVI's within a framework of driver attention support that compares required attention versus supplied attention for a task (Figure 1). If required attention is greater than supplied attention, a countermeasure is needed to increase supplied attention. Countermeasures can also be leveraged when driver behavior, a factor often outside of the control of vehicle designers and manufacturers, falls outside of an acceptable tolerance level. Such situations can occur from internal sources (e.g., choice to use a personal electronic device) or from external sources (e.g., digital billboards, Belyusar et al., 2016). This framework is extendable to both sides of the Yerkes-Dodson curve (Coughlin et al., 2012) to appropriately consider both overload and underload.

| | | | | |
|---|---|---|---|---|
| **Required Attention** | $>$ | **Supplied Attention** | $=$ | **Promote rebuilding of attention using:** Temporary restrictions, Attention prompts and/or other safeguards |
| **Required Attention** | $\leq$ | **Supplied Attention** | $=$ | Enable HMI functionality that follows a principled approach. |

**Fig. 1**. Comparing required and supplied attention.

The framework is explicitly designed to scale across a variety of passenger vehicles equipped with an assortment of sensors and technologies. In this context, required and suppled attention can be inferred to inform different levels of model sensitivity using a range of indirect (e.g., taps) and direct measures (e.g. eye glance data) and driver behaviors. Case studies, shown in Table 1 in Appendix A, demonstrate the scalability of the framework.

### 4.3 Countermeasures

A countermeasure can be defined as an active intervention to realign attention. In the case where supplied visual attention is not sufficient to meet attention required, the aim of an attentional countermeasure is to get the driver looking back at the road. There are three general types of such countermeasures: Adapt, Feedback, and Block.

**Adapt** countermeasures modify the interface's system behavior. For example, adapting, de-cluttering or simplifying displayed information (Chew, et al., 2021), moving drivers to an appropriate interaction modality for the task (*i.e., voice instead of visual-manual*), adjusting or supressing driver feedback at inopportune times (Wright, et al., 2017; Caber, et al., 2023) (and/or suppressing low priority interrupts sent to the driver by subsystems) or even modifying the ADAS system settings to be more sensitive for periods of high demand. **Feedback** countermeasures are active feedback that the vehicle interface uses to nudge the driver to look back at the road. They may also include real-time coaching or brief forms of 'help' or instruction given at carefully selected teachable moments. One example, a real-time prompt that indicates a glance to the road is necessary. Cues could be visual, auditory, haptic, kinaesthetic (*vehicle movement*), pre-attentive or active coaching (*spoken*). Alternatively, the vehicle interface could provide direct and active notification / feedback of threats on the road if a specific threat is of concern. Finally, there is **Blocking** where either dynamically, or permanently, functions are blocked because the situation is too demanding (Leipnitz, et al., 2022). This could take the form of actively stopping a task in progress if the current driving situation requires more attention. Another example would be when the required attention level is such that preventing access to certain tasks, because of the demand that task would place on the driver, could prevent a potential issue.

Countermeasures could and should be used in an escalating fashion if enough foresight can be gained into how quickly the situation could change. Alternatively, if an initial feedback intervention doesn't achieve an increase in supplied attention, then more salient feedback should be triggered. All countermeasures need to be designed carefully, and using a human centred design process to ensure that there is robust evidence that they work to increase supplied attention in an operational environment and do not simply prolong tasks, , nor cause frustration of the driver, nor are too easily ignored.

## 5. Clarification of Scope & Limitations

AHEAD's work to date explicitly acknowledges several **limitations**. Legal requirements need to be maintained until or unless modified. There are a set of in-vehicle activities whose type, nature, and/or demands lie beyond what many feel should be socially acceptable while driving (e.g., watching video). While what is socially acceptable and legally required may evolve over time, it's recognized that in some global markets legally required lockouts need to be respected.

This framework does not identify a specific set of demand limits for what may be considered socially acceptable tasks. OEMs may consider benchmarking this approach to traditional (e.g., radio tuning) or other tasks to ensure that demand considerations meet their organizational philosophies and regulatory commitments. Whatever route is taken, it is important to make sure that limits are credible, evidence-based, and representative of real-world driving.

With regard to scope, one important clarification relates to the use of automated or partially automated driving features by a driver. With the approach described here, the level of assisted or partially automated driving is **viewed as an input into the situationally appropriate attention equation**. In this context, the current framework does not argue that drivers should or should not be provided any additional liberties to engage in secondary tasks under any type of assisted or partially automated driving.

L3 systems dramatically shift the relationship between the driver and vehicle. The framework recognizes that if L3 driving systems are engaged, drivers may be permitted to engage in activities that are not optimized for the driving situation. Future extensions to the framework could encompass elements of L3 operation but, for now, have been considered out of scope.

## 6. Conclusion

While there is still much to be learned about driver behavior with secondary tasks, many of the historical limitations that framed early driver demand guidelines can now be reasonably addressed. This is an **opportune time to build upon the foundations** of prior work and harness new findings and capabilities **to focus on more effective ways** of designing DVIs and related systems to support drivers and mitigate issues such as portable electronic device use.

AHEAD aims to promote these perspectives as an alternate path or approach (not necessarily a replacement) for current guidelines for DVI design, validation, and testing. The premise is to build upon new insights in attention-centric design to, in real-time, assess whether drivers are paying sufficient attention for the current situation and, if not, leveraging technology to support the rebuilding of attention. Where needed, countermeasures can also provide attention triggered failsafe actions.

This framework moves the language of DVI assessment beyond previous efforts to consider: The role of spatial and temporal characteristics of a task; a framework in which demand can be optimized across all dimensions, i.e., visual, auditory, haptic, vocal, manual, etc., by taking into consideration the relative cost and benefit interactions of various input, output and processing modalities, and interactions between secondary tasks and the broader operating environment. As such, assessment moves from focusing narrowly on distraction to a broader consideration of driver attention support and safe operation that emphasises mechanisms that promote rebuilding situation awareness which can:

- Reduce exposure (Seppelt et al., 2018) to unfolding conflicts
- Foster less surprise (Meyer et al, 2022)
- Encourage more measured responses (Seppelt et al., 2017)
- Improve driver readiness
- Result in fewer crashes (Seaman et al., 2017; Seaman 2021) when exposed to a conflict

The development of this work will continue to evolve through the integration of input from interested parties. Efforts to date are explicitly neutral regarding the need for new or updated policy or industry guidelines. We hope that by sharing our work, relevant global organizations can leverage it in their research and that this effort will encourage a broader discussion of next generation driver focus principles and lead to safer, more satisfying travel on the world's roadways.

## 7. Acknowledgments

The authors wish to acknowledge Bobbie Seppelt, Dan Clifton, Code Grebe, Christian Lorenz, Shani Jayant, Elizabeth Wykes, and Pnina Gershon, who all contributed to the development of this work.

**References**

Angell, L.S. (2012). Driver attention & scanning: a focus for enhancing safety in the connected vehicle era. Invited address given to the *Center for Automotive Research (CAR)*, Breakfast Briefing Series. September 28, 2012.

Angell, L. S., Auflick, J., Austria, P. A., Kochhar, D., Tijerina, L., Biever, W., et al. (2006). Driver Workload Metrics Task 2 Final

Report (DOT HS 810 635). Washington, DC: U.S. Department of Transportation National Highway Traffic Safety Administration.

Belyusar, D., Reimer, B., & Mehler, B. (2016). A field study on the effects of digital billboards on glance behavior during highway driving. *Accident Analysis & Prevention, 88,* 88-96.

Bents, F.D. (2000). Driving with cell phones: What have highway safety researchers learned. Comments, Driver Distraction Public Meeting, July 18.

Burnett, G., Neila, N., Crundall, E., Large, D., Lawson, G., Skrypchuk, L., & Thompson, S. (2013). How do you assess the distraction of in-vehicle information systems? A comparison of occlusion, lane change task and medium-fidelity driving simulator methods. *Proceedings of DDI2013*.

Caber, N., Ahmad, B. I., Liang, J., Godsill, S., Bremers, A., Thomas, P., ... & Skrypchuk, L. (2023). Driver profiling and Bayesian workload estimation using naturalistic peripheral detection study data. *IEEE Transactions on Intelligent Vehicles*.

Chew, J. Y., Kawamoto, M., Okuma, T., Yoshida, E., & Kato, N. (2021). Adaptive attention-based human machine interface system for teleoperation of industrial vehicle. *Scientific Reports*, *11*(1), 17284.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, *4*(5), 170-178.

Coughlin, J.F., Reimer, B. & Mehler, B. (2011). Monitoring, managing and motivating driver safety and well-being. *IEEE Pervasive Computing, 10*(3), 14-21.

Ding, L. Terwilliger, J. Parab, A., Wang, M., Fridman, L., Mehler, B., & Reimer, B. (2023). CLERA: A unified model for joint cognitive load estimation and eye region analysis in the wild. *ACM Transactions on Computer-Human Interaction, 30*(6), 1-23.

Hendricks, D. L., Fell, J. C., and Freedman, M. (2001, January). Summary Technical Report: The relative frequency of unsafe driving acts in serious traffic crashes. Washington, DC: National Highway Traffic Safety Administration.

Kircher, K., & Ahlström, C. (2009). Issues related to the driver distraction detection algorithm AttenD. In *First International Conference on Driver Distraction and Inattention. Gothenburg, Sweden*.

Large, D. R., Burnett, G., Crundall, E., Lawson, G., Skrypchuk, L., & Mouzakitis, A. (2019). Evaluating secondary input devices to support an automotive touchscreen HMI: A cross-cultural simulator study conducted in the UK and China. *Applied Ergonomics*, *78*, 184-196.

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.

Mattes, S. (2003). The lane-change-task as a tool for driver distraction evaluation. *Quality of Work and Products in Enterprises of the Future, 57*(60), 7.

Mattes, S., & Hallén, A. (2009). Surrogate distraction measurement techniques: The lane change test. Driver Distraction: Theory, Effects, and Mitigation, 107-121.

Meyer, J., Llaneras, E. & Fitch, G.M. (2022). Impact of Temporary browsing restrictions on drivers' situation awareness when interacting with in-vehicle infotainment systems. *Safety, 8*(81).

Morgenstern, T., Wögerbauer, E. M., Naujoks, F., Krems, J. F., & Keinath, A. (2020). Measuring driver distraction–Evaluation of the box task method as a tool for assessing in-vehicle system demand. *Applied Ergonomics, 88*, 103181.

Rosenholtz, R., Dorai, A., & Freeman, R. (2011). Do predictions of visual perception aid design?. *ACM Transactions on Applied Perception (TAP)*, *8*(2), 1-20.

Scott, D. (1993). Visual search in modern human-computer interfaces. *Behaviour & Information Technology*, *12*(3), 174-189.

Schnelle-Walka, D., & Radomski, S. (2019). Automotive multimodal human-machine interface. In *The Handbook of Multimodal-Multisensor Interfaces: Language Processing, Software, Commercialization, and Emerging Directions-Volume 3* (pp. 477-522).

Seaman, S., Lee, J., Seppelt, B., Angell, L., Mehler, B. & Reimer, B. (2017). It's all in the timing: using the AttenD algorithm to assess texting in the nest naturalistic database. *Proceedings of the 9th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*. Manchester, VT. 403-409.

Seaman, S., Seppelt, B., Angell, L., Reimer, B. & Mehler, B. (2021). Vehicle control and response to emerging events: it's both off-road and on-road glance duration. (extended abstract) *Proceedings of the 7th International Conference on Driver Distraction and Inattention*. 27-31.

Seppelt, B., Seaman, S., Angell, L.S., Mehler, B., & Reimer, B. (2017). Differentiating cognitive load using a modified version of AttenD. *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutoUI*), Oldenburg, Germany, September 24-27.

Seppelt, B., Seaman, S., Angell, L., Mehler, B. & Reimer, B. (2018). Assessing the effect of in-vehicle task interactions on attention management in safety-critical events. *Proceedings of the 6th International Conference on Driver Distraction and Inattention*. Gothenburg, Sweden.

Seppelt, B., Seaman, S., Lee, J., Angell, L.S., Mehler, B. & Reimer, B. (2017). Glass half-full: use of on-road glance metrics to differentiate crashes from near-crashes in the 100-car data. *Accident Analysis and Prevention, 107*, 48-62.

Stutts. J. C., Reinfurt, D. W., Staplin, L., and Rodgman, E. A. (2001, May). The role of driver distraction in crashes. Washington, DC: AAA Foundation for Traffic Safety.

Reimer, B., Angell, L., Mehler, B., Fitch, G.M., Nobel, A.M., Feit, S. & Skrypchuk, J. (2022). Managing distraction holistically through driver attention support. *Proceedings of the 66th Annual Meeting of the Human Factors and Ergonomics Society*. Atlanta, GA, 310-314.

Reimer, B. Angell, L., Strayer, D., Tijerina, L. & Mehler, B. (2016). Evaluating demands associated with the use of voice-based in-vehicle interfaces. *Proceedings of the 60th Annual Meeting of the Human Factors and Ergonomics Society*. Washington, DC. 2083-2087.

Wang, J. S., Knipling, R. R., and Goodman, M. J. (1996). The role of driver inattention in crashes: New statistics from the 1995 Crashworthiness Data System. *Proceedings of the 40th Annual Meeting of the Association for the Advancement of Automotive Medicine* (pp. 377- 392), Vancouver, British Columbia.

Wang, Y., Reimer, B., Dobres, J., & Mehler, B. (2014). The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand. *Transportation Research Part F: Traffic Psychology and Behaviour, 26,* 227-237.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 0058.

Wright, J., Stafford-Fraser, Q., Mahmoud, M., Robinson, P., Dias, E., & Skrypchuk, L. (2017, September). Intelligent scheduling for in-car notifications. In *2017 IEEE 3rd International Forum on Research and Technologies for Society and Industry (RTSI)* (pp. 1-6). IEEE.

## Appendix A

In this framework, each case builds upon the previous one; we begin with rudimentary, indirect measures of driver attention and driving context and progress toward direct measures of glances and behaviors. It is important to note that framework implementation is not necessarily a linear process, manufacturers do not have to begin at Case 1 and they can skip cases if it better aligns with their organization's goals and philosophies for driver assistance.

Changes between cases are bolded so you can easily see what parameters are added.

- Starting with Case 1: we leverage fixed parameters for required attention, like the NHTSA/DFT guidelines, however, we incorporate measures of supplied attention by using interactions with the in-vehicle infotainment system.
- In Case 2 we leverage limited vehicle sensing data such as speed and steering angle to obtain information about the driving environment, while continuing to use IVIS interactions as the supplied attention metric.
- Case 3 is applicable to vehicles with additional sensing capabilities, such as radar, which provides even more information about the driving environment.
- Case 4 keeps the same parameters for required attention, but adds indirect measures of supplied attention, such as steering entropy.
- Case 5 adds moment to moment driving risks to the required attention assessment – this can include longitudinal conflicts, lateral conflicts, lane departures, and more.
- Case 6 adds in direct driver monitoring measures – that could include driver glance metrics such as eyes off road time or glances to specific areas of interest.
- Case 7 adds in parameters of behavioral modelling – going beyond glance metrics and incorporating measures of driver behavior and workload such as non-driving related tasks, drowsiness, and fatigue.

As this set of cases shows, this is a scalable solution, which allows this framework to be applied across wide range of vehicles without mandating any additional technologies.

*Table 1. Driver Attention Support Framework Case Studies*

| Case | Key Factors for Required Attention | Key Factors for Supplied Attention |
|---|---|---|
| **Case 1: Leveraging supplied attention alone** | Fixed | Using interactions with in-vehicle information systems (IVIS) |
| **Case 2: Context dependent levels of required attention** | **Using limited vehicle sensing (e.g., current speed, steering angle)** | Using interactions with IVIS |
| **Case 3: Enhanced context dependent levels of required attention** | **Using additional vehicle sensing (e.g., current speed, steering angle, ACC/TTC)** | Using interactions with IVIS |
| **Case 4: Enhanced assessment of supplied attention using indirect measures of supplied attention** | Using vehicle sensing (e.g., current speed, steering angle, ACC/TTC) | Using interaction with IVIS and **supplemented/supported by indirect measures of driver attention (e.g., steering entropy)** |
| **Case 5: Extending to moment-to-moment driving risks** | Using vehicle sensing (e.g., current speed, steering angle, ACC/TTC) and **moment-to-moment driving risks** | Using interaction with IVIS and supplemented/supported by indirect measures of driver attention (e.g., steering entropy) |
| **Case 6: Incorporating direct measurement of driver attention** | Using vehicle sensing (e.g., current speed, steering angle, ACC/TTC) and moment-to-moment driving risks | Using interactions with IVIS **and direct driver attention monitoring** |
| **Case 7: Driver state monitoring outputs as supplied attention modifiers** | Using vehicle sensing (e.g., current speed, steering angle, ACC/TTC) and moment-to-moment driving risks | Using interactions with IVIS, direct driver attention monitoring and **other behavior monitoring** |

# Implementation Guide to the Australian Driver Distraction Roadmap

Matin Nabavi Niaki [1*], Martin Small [2], Michael Regan[3]

[1] Austroads, Road Safety and Design, Level 17, 360 Elizabeth St., Melbourne, Australia
[2] Martin Small Consulting Pty Ltd, Adelaide, Australia
[3] Research Centre for Integrated Transport Innovation, University of NSW, Sydney, Australia
[*]MNabavi@austroads.com.au

**Abstract: This paper describes the development of an Implementation Guide prepared in response to a commitment in Australia's National Road Safety Strategy 2021-2030 to implement a National Driver Distraction Roadmap. This guide addresses driver distraction comprehensively, moving beyond behavioural interventions to recognize its systemic nature within the road traffic system. It emphasizes a need for understanding the complex interplay of factors involved. The guide outlines a structured methodology, including problem definition, current state assessment, forward work program development, and engagement with various stakeholders. By advocating for collaboration, innovation, and ongoing monitoring, the guide offers a holistic approach to prevent distraction-related crashes and injuries.**

## 1. Introduction

Driver distraction has traditionally been approached as a behavioural issue in road safety, prompting initiatives like awareness campaigns and law enforcement targeting mobile phone use. However, it is recognized as a complex phenomenon stemming from interactions within the road traffic system, including vehicle design, user behaviour, and environmental factors such as the design of roadside infrastructure. Social and psychological influences, such as mental health and social media, further compound the issue.

While current approaches emphasize the need to accommodate human error through a more forgiving road environment, more supportive active and passive vehicle technology and better speed management, these will not fully address distraction.

The Implementation Guide described here seeks to prevent distraction-related errors by motor vehicle drivers through a systemic approach. Emphasizing design solutions and a deeper understanding of system interactions, it aims to move beyond accommodating distracted driving to effectively mitigating distraction risks at their root.

### *1.1 Background*

Assessing the scale of the distraction issue proves challenging due to limited data capturing specific distractions. Crash reports often lack evidence of mobile phone use or other distractions. An Australian study found distractions in 45% of casualty crashes (Fitzharris et al. 2022), yet global analyses warn of potential over-/under-estimation of the problem due to insufficient behavioural detail and variability in definitions and coding of distraction (Regan and Oviedo-Trespalacios 2022; IIHS 2022).

While a wide range of potential distractions for a driver or rider had been present for decades, the rising ubiquity of mobile phones led to them being recognised as a major new injury risk. Research has subsequently shown that dialling, texting, and talking on a mobile phone while driving can lead to riskier decision making, slower reactions, speed and vehicle control variations, and less controlled braking – the driver will tend to brake later, with more force and less control (Oviedo-Trespalacios 2018; Li et al. 2020; Haque and Washington 2014A; Oviedo-Trespalacios et al. 2018; Haque and Washington 2014B).

### *1.2 Current Study in Context*

While existing countermeasures have been to some extent effective in eliminating/reducing the occurrence/severity of distraction-related crashes, the Implementation Guide described here adopts a system perspective on distraction-specific activities. This involves the proposal of a structure that systematically encompasses all aspects and stakeholders of distracted driving.

## 2. Method

Development of an implementation guide from a systems perspective encompasses all aspects related to distracted driving. Table 1 and the following subsections outlines the proposed stages of the Implementation Guide.

### *2.1 The Problem*

Distraction can cause unsafe responses from a driver, in four primary ways taking attention away from activities critical for safe driving (CARRSQ 2020):

- eyes off the road caused by visual distraction,
- mind off the road caused by cognitive distraction,
- ears off the road caused by cognitive distraction,
- hands off the controls caused by physical interference.

### *2.2 Current State*

The driver distraction problem should be defined from a systems perspective. This step requires the familiarisation with and knowledge of current and future driver distraction *development programs*. These development programs include vehicle technologies such as Advanced Driver Distraction Warning (ADDW), which is scheduled for implementation in all new vehicles from 2029 in the European Union. Technology neutral *legislation* on portable device use (Wolter, 2022), as well as *enforcement programs* such as automatic mobile phone use detection from roadside cameras, are also important, as is research on the design of in-vehicle communications systems. Table 1 below outlines the proposed stages of the Implementation Guide and the approach to be taken to tackle distracted driving.

***Table 1*** *Implementation guide and approach*

| This Guide | Proposed Approach |
|---|---|
| 1. The Problem | **Orient:** Consider the core issues, and how your organisation can make a meaningful impact on the driver distraction problem |
| 2. Current State | **Familiarise:** Familiarise yourself with the major development programs that are underway and the potential impact they can have |
| 3. Forward Work Program | **Align:** Systematically review the Forward Work Program agreed by stakeholders and how your organisation's effort can align with one or more elements |
| 4. Implementation Partners | **Engage:** Discuss the issue with potential partners including how your interests and responsibilities in distraction prevention coincide and present opportunities |
| 5. Implementation Mechanisms | **Design:** Identify the best means of achieving the change that you want to see, in line with good practice injury prevention principles and current state issues |
| 6. Ongoing Implementation | **Deliver:** Develop and implement systems that will sustain this activity over an extended period, recognising the seemingly intractable nature of the driver distraction problem. Ongoing monitoring and evaluation of the systems to learn from and improve upon |
| 7. Good Practice | **Learn:** Retain a focus on learning from and collaborating with others working in the field. This includes ongoing monitoring and evaluation of programs and countermeasures |

### *2.3 Forward Work Program*

The development of a Forward Work Program agreed by stakeholders and all concerned organisations is a major step in addressing driver distraction. The implementation domains include: design, mitigation, workplace, compliance, behaviour, systems, communications, engagement, and evaluation. Table 2 compiles existing knowledge to outline the forward work program domains and options (Regan and Oviedo-Trespalacios 2022; Department of Transport and Main Roads 2020; Regan et al. 2009; European Commission 2015; PIARC 2016; Imberger et al. 2020; Regan et al. 2020).

### *2.4 Implementation Partners*

There are various participants who have a role to play in preventing driver distraction described below.

#### *2.4.1 Governments and Regulators*

The primary authority, responsibility and power associated with tackling driver distraction lies with government comprised of *parliament*, *executive*, and *judiciary*. Whether at a national, state, or local level, the responsibilities of the elected representatives and decision makers regarding driver distraction are to:

- understand the safety impact of driver distraction,
- respond positively to evidence-based advice on different mitigation strategies,
- make the necessary changes to law, funding, and programs to support driver distraction prevention,
- invest in ongoing monitoring and evaluation of these programs.

#### *2.4.2 Industry/Advocacy Associations*

Local and global motor vehicle manufacturers and their agents must address the human-machine interface that may lead to distraction. Private sector interests in shaping the potential for driver distraction prevention also include:

- Tier 2 suppliers of in-vehicle technology systems,
- producers of the devices which distract drivers,
- suppliers of telecommunications services which facilitate in-vehicle electronic distraction.

#### *2.4.3 Product Developers and Service Providers*

Organisations which develop and sell products and services that are directly connected to driver distraction are:

- motor vehicle manufacturers,
- telecommunications providers,
- producers of goods which are known to be consumed in vehicles, such as fast food, make-up,
- producers of advertising services, for electronic transmission into vehicles, and for the roadside.

As with all participants in the driver distraction ecosystem, these developers and providers have a responsibility to prevent driver distraction within their own operations, ensuring that their products and services are safe.

#### *2.4.4 Corporations and Organisations*

There are many organisations which may have an interest in distraction, either now or in the future, largely due to their work health and safety responsibilities, such as:

- organisations which have a direct link to the transport industry, through the carriage of passengers or goods,
- organisations without a link to the transport industry but have staff and contractors who travel in motor vehicles as a secondary activity.

These organisations providing transport services, or contracting transport services, or which have staff engaged in significant road travel, have an obligation to consider exposure to these safety hazards and the extent to which risks associated with distraction are being managed.

#### *2.4.5 Users*

Distraction prevention has traditionally targeted individual user behaviour, but there is a growing acknowledgment of the inherently distracting environment. While drivers must prioritize attention to driving, manufacturers and designers of vehicles and devices play a crucial role and should implement processes to gather feedback from drivers on distraction-impacted behaviours, enhancing future prevention measures.

**Table 2** *Forward work program*

| Domains | Options |
|---|---|
| **Design:** develop design principles, guidelines, and standards to facilitate safer interactions between drivers, technology, vehicles, and roads. | Develop and implement guidelines and standards for the design of the vehicle human-machine interface (HMI) to reduce distraction. |
| | Adopt Human Factors Integration (HFI) processes to ensure that products and systems are user-centred designed to prevent and mitigate distraction. |
| | Promote standard design of mobile/wearable device to restrict distraction while driving. |
| | Develop and implement guidelines/standards for road/traffic design to reduce distraction. |
| | Develop assessment protocols for rating the existing road and traffic environment for its potenti to distract drivers that could be incorporated in road safety audits and safety star ratings. |
| | Develop standardised criteria/methods for evaluating the safety impact of advertising signage. |
| **Mitigation:** increase the availability and implementation of in-vehicle distraction mitigation technology | Identify potential vendors/suppliers to develop and implement after-market distraction mitigating technologies which could be cost-effectively retrofitted into existing vehicles. |
| | Develop assessment protocols for rating vehicles on their potential to distract drivers and incorporate it in new car assessments to encourage improved human-machine interface. |
| | Promote standardisation of interfaces for the secure placement, mounting and powering of nomadic devices on vehicle dashboards to prevent distraction induced by sliding/ falling. |
| | Stimulate demand for other technologies (such as phone blocking, distraction warning systems) that have proven to prevent and directly mitigate the effects of distraction. |
| **Workplace:** work with employers and workplace health and safety regulators to improve approaches to driver distraction | Investigate the relationship between job demands, wellbeing, and distracted driving. |
| | Encourage employers to develop/implement best practice guidance for managing distraction. |
| | Investigate the use of financial and non-financial incentives on corporate fleet insurance policies through implementation of driver distraction prevention technologies. |
| **Compliance:** strengthen compliance mechanisms through improved rules, detection, and evaluation | Evaluate the effectiveness of existing distraction regulations and penalties. |
| | Investigate the effectiveness of automatic mobile phone detection technologies. |
| | Monitor and trial new technologies that support compliance with distraction regulations. |
| **Behaviour:** shift driver behaviour through innovative campaigns and educational strategies. | Strengthen distraction management education and training for drivers through driver licensing and road safety education processes. |
| | Promote motor vehicle buyers to prioritise technology features that minimise distraction. |
| | Investigate the use of personalised insurance incentives for individuals that exhibit safe driving habits, weighted towards distraction mitigation. |
| **Systems:** develop systematic and ongoing response to engage all parties | Develop a communications platform with a dashboard to support ongoing communications, engagement, and implementation. |
| | Evaluate delivery of this work program after six years and develop a distraction prevention strategy and plan for 2040. |
| **Communications:** implement good practice communications. | Develop a shared national narrative for driver distraction and align industry and manufacturer led educational campaigns to drive cultural change and awareness of distracted driving. |
| **Engagement:** implement engagement methods that bring parties together | Establish and operationalise an ongoing stakeholder-oriented Governance Framework for preventing driver distraction. |
| | Recognise the role that non-transport stakeholders such as the healthcare system or the food and entertainment industries have in distracted driving. |
| **Evaluation:** nourish an ongoing effort to learn more about the issue and what needs to be done | Develop and adopt a common national definition of distraction that can be operationalised and used to code crash and incident data. |
| | Standardise the way distraction data are collected and coded in crash and incident databases. |
| | Provide training for Police and crash investigators to detect distraction as a contributing factor in crashes and distinguish it from other mechanisms of inattention. |
| | Evaluate the effectiveness of all distraction countermeasures. |
| | Undertake Coronial based no-blame crash investigations of distraction-related crashes for formal reporting to responsible Ministers. |
| | Develop a data platform to enable the investigation, tracking, and sharing of crash and infringement data resulting from driver distraction. |
| | Investigate the impact of Advanced Driver Assistance Systems and Partially Automated Driving Systems on driver distraction. |

### 2.5 Implementation Mechanisms

The implementation mechanisms describe the best means of effectively tackling the issue of distracted driving. This will require a new approach to *governance and leadership* which reflects the roles of different stakeholders.

The consumer demand for safe vehicles can be directed towards vehicle manufacturers or infrastructure providers. Therefore, it is important for the *consumer market* to be involved in the process.

There are limits to the ability of *regulation* to improve safety. However, regulation has proven time and again to be an effective means of significantly improving road safety.

*Promotion* is critical, as well as developing and maintaining momentum in preventing driver distraction by linking different elements together into a single narrative.

Effective solutions to the driver distraction problem will only come through *research and development*: innovation, testing, deployment, evaluation, and improvement. Better understanding of driver and societal acceptance of the issue is also important.

### 2.6 Ongoing Implementation

Commitment to ongoing implementation of a distraction prevention agenda will require a forward perspective on the issue and sustained effort. This will be done through a *systems perspective*, effective *stakeholder engagement*, *learning, sharing, and innovating*, as well as *monitoring and evaluation*.

### 2.7 Good Practice

Learning from existing programs, implementations, research, and evaluation through collaborating with others working in the field and sharing ideas and information will significantly add to the progress in effectively addressing driver distraction and its unwanted outcomes.

## 3. Conclusions

This Australian Implementation Guide offers a holistic approach to addressing driver distraction, recognizing its systemic nature within the road traffic system. By emphasizing design-based solutions and engaging stakeholders, it aims to prevent distraction-related crashes and injuries.

Despite challenges in assessing the scale of the issue, the guide advocates for collaboration, innovation, and ongoing monitoring and evaluation to create a safer road environment. Through a systemic perspective and effective implementation mechanisms, it seeks to address distraction at its root causes, contributing to improved road safety and the mitigation of distraction-related risks.

## 4. Acknowledgments

The authors would like to acknowledge Queensland and Australian Governments who produced the National Driver Distraction Roadmap in 2020.

## References

CARRS-Q (2020). *Fact Sheet Mobile phone use & distraction.* Centre for Accident Research & Road Safety – Queensland.

Department of Transport and Main Roads. (2020b). National roadmap on driver distraction. Brisbane: DTMR. Retrievable from https://www.tmr.qld.gov.au.

European Commission. (2015). *Study on good practice for reducing road safety risks caused by road user distractions – Final report*. Brussels: European Commission.

Fitzharris, M.P., Pok Arundell, T., Corben, B., Lenné, M.G., Liu, S., Peiris, S., Stephens, A., Bowman, D.M., Morris, A., & Tingvall, C. (2022). Identification of countermeasures to address serious injury crashes, ECIS Report 4. Monash University Accident Research Centre, Report 346; Clayton.

Haque, M., & Washington, S. (2014A). A parametric duration model on reaction times of drivers distracted by mobile phone conversations. *Accident Analysis & Prevention,* 62, 42-53.

Haque, M., & Washington, S. (2014B). The impact of mobile phone distraction on the braking behaviour of young drivers: A hazard-based duration model. *Transportation Research Part C: Emerging Technologies*, 50, 13-27.

Imberger, K., Poulter, C., Regan, M.A., Cunningham, M.L., & Paine, M. (2020). Considerations for the development of a driver distraction safety rating system for new vehicles. *Journal of Road Safety*, 31(2), 23–34.

Insurance Institute of Highway Safety (2022). *Distracted Driving*. Arlington. Accessed January 2023. https://www.iihs.org/topics/distracted-driving

Li, X., Oviedo-Trespalacios, O., & Rakotonirainy, A. (2020). Drivers' gap acceptance behaviours at intersections: A driving simulator study to understand the impact of mobile phone visual-manual interactions. *Accident Analysis & Prevention,* 138, 105486.

Oviedo-Trespalacios, O. (2018). Getting away with texting: Behavioural adaptation of drivers engaging in visual-manual tasks while driving. *Transportation Research Part A: Policy and Practice*, 116, 112-121.

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.

PIARC. (2016). *The role of road engineering in combatting driver distraction and fatigue road safety risks*. Report 2016R24EN. Paris: World Road Association (PIARC).

Regan, M.A., & Oviedo-Trespalacios, O. (2022). Driver Distraction: Mechanisms, Evidence, Prevention, and Mitigation, in K.E. Björnberg et al. (eds.) (2022), *The Vision Zero Handbook: Theory, Technology and Management for a Zero Casualty Policy*.

Regan, M.A., Lee, J.D., & Young, K. (2009). Driver distraction: Theory, effects and mitigation. Boca Raton, FL: CRC Press.

Regan, M.A., Prabhakharan, P., Wallace, P., Cunningham, M.L., & Bennett, J.M. (2020). *Education and training for drivers of assisted and automated vehicles*. Austroads research report AP-R616-20. Sydney: Austroads.

Wolter, J. (2022). Managing the driver distraction risk posed by technology: Changes to Australian road laws aimed at driver behaviour. *Proceedings of the 8th International Conference on Driver Distraction and Inattention*, Gothenburg, Sweden.

# Improving take-over performance during cognitive distraction with an eye-gaze and situation adaptive human-machine interface

Gaëtan Merlhiot [1*]

[1] Human Factors, VEDECOM Institute, mobiLAB - 23 bis allée des Marronniers, 78000 Versailles, FRANCE
[*] gaetan.merlhiot@vedecom.fr

**Abstract: Automated driving (AD) (i.e., SAE level 3 AD) allows drivers to do something else, yet they can be required to takeover when the AD ends. We tested how an eye-gaze and situation adaptive human-machine interface (HMI) could improve take-over performance when drivers could be cognitively distracted by a conversation. We hypothesized that the eye-gaze and situation adaptive HMI would improve the reconstruction of situation awareness (SA) (the fact of understanding a situation and being able to anticipate the next events) and thereby take-over performance. We conducted a driving simulator experiment with 42 participants (14 per HMI condition: control with a basic HMI, situation adaptive HMI, eye-gaze and situation adaptive HMI) who drove on the highway with a SAE level 3 AD activated and experienced take-over requests. The eye-gaze and situation adaptive HMI improved: (i) SA reconstruction since participants presented higher gaze dispersion, faster RT to the left mirror and spent more time looking at the left mirror, which indicated more visual searches and more information gathered during the take-overs; (ii) take-over performance and safety with increased time to collision and fewer collisions. Our HMI therefore exhibited promising results. Overall, adaptive HMIs seem to be a good option to ensure better take-over performance safety compared to non-adaptive HMIs.**

## 1. Introduction

Driving automation provides the driver different level of support, from an automated driving (AD) assistance in a SAE level 2 to an AD feature in a SAE level 3 (SAE international, 2016). With higher levels of automation, the driver is progressively moved away from the driving loop, resulting in a decreased situation awareness (SA), namely, the fact of understanding a situation and being able to anticipate the next events (Endsley, 1995). Moreover, the AD can end leading the driver to takeover which requires quickly regaining the SA (Morales-Alvarez et al., 2020).

The period during which the driver regains SA are cognitively demanding (i.e., requiring cognitive and visual attention, see Gold et al., 2015; Kim et al., 2018) and can be impacted if the driver was distracted by a non-driving related task (NDRT), such as having a phone conversation (for a review, see Merlhiot & Bueno, 2021; Zhang et al., 2019).

As mentioned, SA reconstruction relies on visual search that depends on two independent visual attentional processes (e.g., Connor et al., 2004; Katsuki & Constantinidis, 2014; Pinto et al., 2013; Theeuwes, 2010). (i) A bottom-up process that arises from the salience of a stimuli, based on its characteristics in a visual scene. (ii) A top-down process that corresponds to the driver's attention to visual elements that can be guided by knowledge or feedback. These processes can be impacted by cognitive load (i.e., the amount of cognitive resource used), which reduces visual attention from top-down processes and, conversely, increases visual attention from bottom-up processes (Longstaffe et al., 2014). We thus considered assisting the reconstruction of SA by highlighting important visual elements (bottom-up effect) and provide feedback regarding the driver's visual search (top-down effect) by using an eye-gaze and situation adaptive human-machine interface (HMI).

We hypothesized that the eye-gaze and situation adaptive HMI would improve the reconstruction of SA and thereby take-over performance by supporting the two processes associated to a visual search as follow: (i) A bottom-up effect was achieved by using visual cues (i.e., with augmented reality) on important information to direct visual attention to relevant information for the SA reconstruction. (ii) A top-down effect was obtained by giving auditory feedback to focus attention on any important information related to the SA reconstruction that the driver may have missed.

## 2. Method

### 2.1 Participants

We obtained a sample of 42 participants (23 females, 19 males) with a mean age of 35.64 ± 9.78 years. All participants signed a written informed consent form and received 40 euros for their participation.

### 2.2 Driving simulator and eye-tracker

We used a medium-fidelity static driving simulator, which ran under SCANeR studio 2023.1 (Avsimulation®) with a synchronized eye-tracker (Smart Eye Pro 6.0) sampled at 60 Hz.

### 2.3 Experimental design

The study was conducted in a driving simulator with a car equipped with a SAE level 3 highway AD. During the AD, participants experienced take-over requests (TOR) which involved lane changes with ongoing traffic while being cognitive distracted by the twenty-questions task (TQT) (Merat et al., 2012), which simulates a conversation to get closer to ecological settings of AD.

We used the properties of the adaptive HMI as a between-subject variable to created three groups of participants (control with a basic HMI, situation adaptive HMI, eye-gaze and situation adaptive HMI). Our design included the TOR time-budget as a within-subject variable:

urgent (i.e., TOR = 7 seconds) and planned TOR (i.e., TOR = 30 seconds).

*2.3.1 TOR time-budget and use-cases:*

We used 3 urgent and 2 planned TORs, all occurred on the highway at 130 km/h with the ego-vehicle in the right lane and with surrounding traffic. The first four TORs required a lane change due to an obstacle (i.e., stopped vehicle, work zone, lane reduction) and the last one required an emergency braking.

*2.3.2 Eye-gaze and situation adaptive HMI:*

The situation adaptive part of the HMI highlighted important elements and used a 3-colors code based on TTC (Fig. 1): red (TTC < 2s) ; yellow (TTC = [2s,4s]); green (TTC > 4s). The eye-gaze adaptive part of the HMI corresponded to auditory feedback (e.g., check the rear mirror) that was emitted after a certain time (i.e., 2s after Urgent TOR, 10s after Planned TOR) if an area of interest (i.e., road, rear mirror, left mirror) was missed.

***Fig. 1.*** *Example of the situation adaptive HMI*

### *2.4 Procedure*

Participants started with a 15-minutes training session during which they familiarized with the driving simulator and experienced a TOR. After that, they completed the main scenario that lasted about 30 minutes.

### *2.5 Statistical analysis*

All analysis were obtained from linear mixed models with the two independent variables as fixed factors and participants as random variable. Pairwise comparisons used Bonferroni corrections.

## 3. Results

### *3.1 Eye-gaze metrics*

In the *"Eye-gaze and situation adaptive HMI"* condition, participants exhibited higher gaze dispersion, $F(2,40.5) = 3.74$, $p = .011$, faster reaction time (RT) to the left mirror, $F(2,191) = 6.91$, $p < .001$, and higher percentage of time spent looking at the left mirror, $F(2,43.9) = 4.57$, $p = .016$, (Fig. 2).

We obtained an interaction effect between the HMI and TOR time-budget conditions for the percentage of participants who checked at least one mirror (i.e., left mirror, rear mirror) before performing an action (i.e., changing lane, braking), $F(2,154)=3.51$, $p = .032$, (Fig. 3).

In the *"Eye-gaze and situation adaptive HMI"* condition, participants received feedback to check the rear mirror 90% of the time, 95% CI [83%,97%], and to check the left mirror in 61% of the time, 95% CI [50%,73%]. After receiving feedback, participants checked within 3 seconds the rear mirror 15% of the time, 95% CI [3%,27%], and the left mirror 80% of the time, 95% CI [67%,93%], which was different from randomness, $F(1, 62)=54.1$, $p < .001$, $F(1, 42)=27.3$, $p < .001$, respectively.

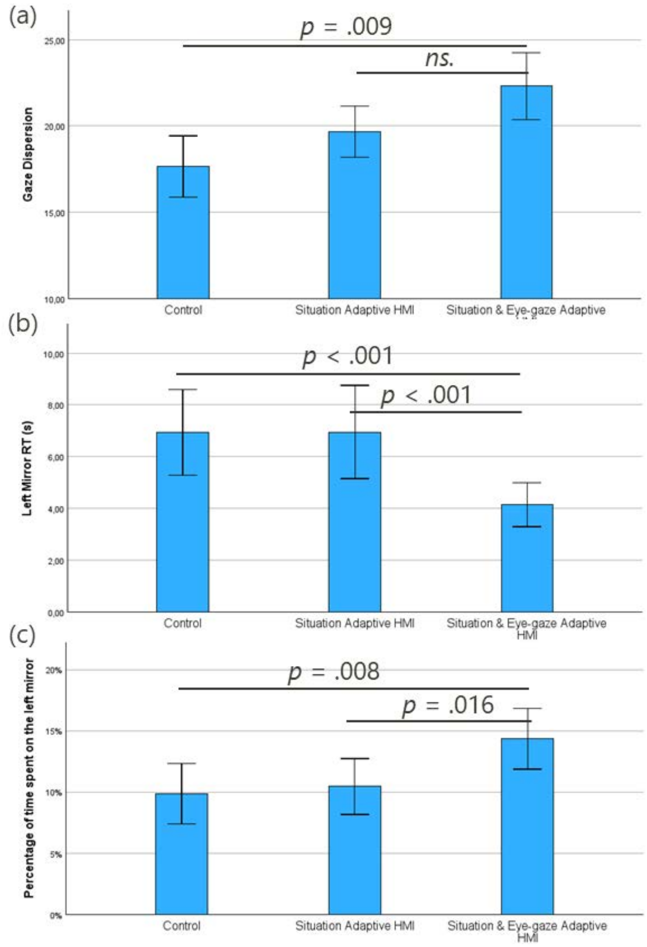



***Fig. 2.*** *(a) Gaze dispersion, (b) left mirror RT and (c) percentage of time spent looking at the left mirror*

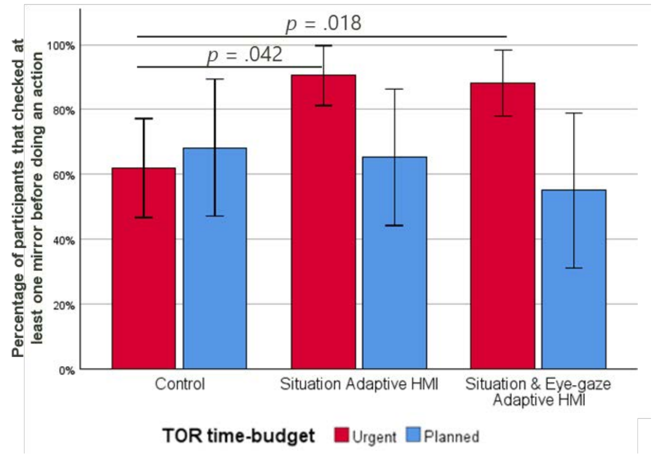



***Fig. 3.*** *Interaction effect between HMI and TOR time-budget conditions for the percentage of participants who checked at least one mirror before performing an action.*

### 3.2 Driving behavior metrics

Participants in the *"Eye-gaze and situation adaptive HMI"* condition presented increased TTC and reduced collision rate in comparison to the other conditions, $F(2,40) = 5.08$, $p = .011$ and $F(2,204) = 9.46$, $p < .001$, respectively (Fig. 4).

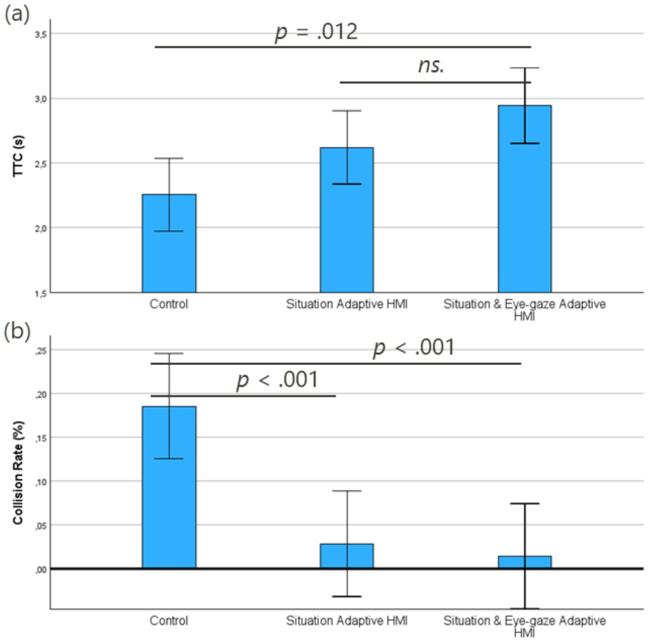



***Fig. 4.*** *(a) TTC and (b) collision rate in function of the HMI conditions*

## 4. Discussion

The aim of this study was to test the effectiveness of an eye-gaze and situation adaptive HMI to improve SA reconstruction and thereby take-over performance and safety in more ecological settings, such as with a variety of use-cases and the driver being cognitively distracted. According to our results, the eye-gaze and situation adaptive HMI improved: (i) SA reconstruction since participants presented higher gaze dispersion, faster RT to the left mirror and spent more time looking at the left mirror, which indicated more visual searches and more information gathered during the take-overs; (ii) take-over performance and safety with increased time to collision and fewer collisions.

Regarding limitations: In some cases, the differences between the two adaptive HMI conditions were not significant, which could arise from a lack of statistical power. We did not use a full factorial design; thus, we did not include a simple eye-gaze adaptive HMI due to time and sample constraints. We focused on the whole HMI, still it would be very interesting to add this condition for further analysis.

## 5. Conclusions

Our eye-gaze and situation adaptive HMI exhibited promising results with improved SA reconstruction and increased take-over performance and safety. Further analysis will focus on eye-gaze data with Markov analysis and physiological data. Overall, the adaptive HMIs seem to be a good option to ensure better take-over performance safety in comparison to non-adaptive HMIs.

## 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. Neither the European Commission nor CINEA – in its capacity of Granting Authority – can be made responsible for any use that may be made of the information this document contains.

## References

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: Bottom-up versus top-down. *Current Biology*, *14*(19), 850–852. https://doi.org/10.1016/j.cub.2004.09.041

Endsley, M. R. (1995). Article in Human Factors The Journal of the Human Factors and. In *Human Factors Journal* (Vol. 37, Issue 1). Ergonomics Society. https://www.researchgate.net/publication/21019B492

Gold, C., Berisha, I., & Bengler, K. (2015). Utilization of drivetime - Performing non-driving related tasks while driving highly automated. *Proceedings of the Human Factors and Ergonomics Society*, *59*(1), 1666–1670. https://doi.org/10.1177/1541931215591360

Katsuki, F., & Constantinidis, C. (2014). Bottom-up and top-down attention: Different processes and overlapping neural systems. *Neuroscientist*, *20*(5), 509–521. https://doi.org/10.1177/1073858413514136

Kim, J., Kim, H. S., Kim, W., & Yoon, D. (2018). Take-over performance analysis depending on the drivers' non-driving secondary tasks in automated vehicles. *9th International Conference on Information and Communication Technology Convergence: ICT Convergence Powered by Smart Intelligence, ICTC 2018*, 1364–1366. https://doi.org/10.1109/ICTC.2018.8539431

Longstaffe, K. A., Hood, B. M., & Gilchrist, I. D. (2014). The influence of cognitive load on spatial search performance. *Attention, Perception, and Psychophysics*, *76*(1), 49–63. https://doi.org/10.3758/s13414-013-0575-1

Merat, N., Jamson, A. H., Lai, F. C. H., & Carsten, O. (2012). Highly automated driving, secondary task performance, and driver state. *Human Factors*, *54*(5), 762–771. https://doi.org/10.1177/0018720812442087

Merlhiot, G., & Bueno, M. (2021). How drowsiness and distraction can interfere with take-over performance: A systematic and meta-analysis review. *Accident Analysis & Prevention*, *December*, 106536. https://doi.org/10.1016/j.aap.2021.106536

Morales-Alvarez, W., Sipele, O., Léberon, R., Tadjine, H. H., & Olaverri-Monreal, C. (2020). Automated driving: A literature review of the take over request in conditional automation. *Electronics (Switzerland)*, *9*(12), 1–34. https://doi.org/10.3390/electronics9122087

Pinto, Y., van der Leij, A. R., Sligte, I. G., Lamme, V. A. F., & Scholte, H. S. (2013). Bottom-up and top-down attention are independent. *Journal of Vision*, *13*(3), 16. https://doi.org/10.1167/13.3.16

SAE international. (2016). Automated Driving - Levels of Driving Automation Are Defined in New Sae

International Standard J3016. *SAE International*. https://doi.org/P141661
Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006
Zhang, B., de Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *64*, 285–307. https://doi.org/10.1016/j.trf.2019.04.020

# How Technology Can Help Reduce Driver Distraction

Pnina Gershon[1*], David Zuby [2], Ian Reagan[2], Greg Fitch[3], Scott Schmidt[4], Mike Lenne[5] , and Mike Hernandez[4]

[1] Center for Transportation & Logistics, AgeLab, MIT, 1 Amherst Street, Cambridge, MA, USA
[2] Insurance Institute for Highway Safety, 4121 Wilson Blvd. Arlington, VA, USA
[3] Google, 1600 Amphitheatre Pkwy, Mountain View, CA, USA
[4] Alliance for Automotive Innovation, 1050 K Street, NW, Washington, DC, USA
[5] Seeing Machines: 271 Collins Street, Melbourne, VIC, Australia
[*]pgershon@mit.edu

**Abstract:** Distracted driving is a leading contributing factor to vehicle crashes in many countries around the world. Minimizing driver distraction is particularly relevant given the technological evolution of infotainment and other in-vehicle systems. This paper reviews existing technologies and strategies to reduce driver distraction, encompassing monitoring-based prevention, restriction-based prevention, utilization of crash avoidance features, and automated enforcement. While these approaches hold promise, further research is required to evaluate their effectiveness and implementation. Additionally, regulatory frameworks play a crucial role in addressing driver distraction, with guidelines and legislation shaping the adoption of distraction-mitigation technologies. Understanding human-technology interaction and addressing regulatory disparities are essential for the successful integration of distraction-prevention measures. Ultimately, advancements in technology and policy are pivotal in counteracting distracted driving and enhancing road safety.

## 1. Introduction

Distracted driving is a specific case of inattention when non-driving tasks capture the driver's attention such that the driver becomes oblivious to the road and traffic events [1]. In 2020, over 3,000 people in the U.S. were killed and over 300,000 were injured in distraction-related crashes [2]. Distracted driving and resulting crash risk may vary depending on many factors related to the type of task, driver characteristics, driving environment, and the vehicle where driving automation and active safety systems are assisting the driver and reducing workload [3, 4]. Nowadays, consumer vehicles are increasingly equipped with partial-automation systems that can simultaneously control the longitudinal and lateral vehicle kinematics on a sustained basis. The driver still remains responsible for monitoring, object/event detection, response selection, and execution. As the driver role pivots toward monitoring and the driving demands are lowered, it is very difficult to maintain attention to the driving task and we see evidence of an increase in undesired behaviors related to distracted driving [5]. Drivers often use the "freed-up" resources to do other things and this tendency is amplified even more by the increased availability of portable electronics and in-vehicle technologies in the form of entertainment, navigation, information, and communication systems. As such, minimizing driver distraction is particularly relevant.

In parallel, emerging technological advancements can also offer new ways to mitigate driver distraction. The opportunity to develop technology-based countermeasures for distracted driving as part of the development of driving automation calls for policies and regulations that will foster innovation and guide implementation without restricting progression. These approaches should be an integral part of the automation and be built from the ground up. This work summarizes existing approaches and technologies to prevent and/or reduce the negative consequences of driver distraction. This work also identifies gaps in current technology and outlines where future work can contribute to prevention.

## 2. Technologies to Mitigate Distracted Driving

### *2.1 Monitoring-Based Prevention*

Driver-monitoring systems (DMS) typically alert inattentive drivers to look back at the road. DMS with indirect driver monitoring infers driver state using vehicle control measures (e.g., steering or throttle inputs), driving duration, and other inputs. Direct driver monitoring relies on camera-based methods, which affords more, specificity in identifying driver states and risky behaviors. More advanced approaches use combinations of metrics to classify when a driver is disengaged from the driving task. To achieve the desired crash reductions, it is important that the system not only detect and warn of an impaired driver state, such as distraction but that the driver state is also communicated to the other safety systems in the vehicle. For example, increasing the sensitivity of driver assistance systems when a driver is classified as not attentive. The next steps call for studying the efficacy and effectiveness of DMS in reducing risks associated with distraction and evaluating different strategies that combine warnings and interventions. This way the potential distractions can be reduced to match both the requirements imposed by the driving environment and driver attention allocation.

### *2.2 Restriction-Based Prevention*

Drivers may benefit from technologies that limit the opportunity to engage in distracting behaviors. Ford's MyKey is one example of an in-vehicle technology suite that includes features like speed control and features that block incoming text and calls while the vehicle is in motion. Smartphone-based blocking technologies, services, or applications, can limit distracted driving by prohibiting calls and texts, blocking audio features, and specific apps.

Cellphone manufacturers also offer a Do Not Disturb Driving mode (DNDD) that can help drivers stay focused on the road by silencing or limiting notifications when driving. DNDD mode permits voice interaction and allows messages from select contacts to be automatically read out. Research to understand the effectiveness of DNDD apps in reducing crashes related to cellphone distraction would be valuable. Other research gaps center on driver awareness and acceptance of DNDD apps and on ways to improve messaging on their availability and evolving functionality. The emergence of telematics as a data source presents an opportunity for such efforts.

### *2.3 Using Crash Avoidance Features for Prevention*

While not designed specifically to address driver distraction, crash avoidance features can alert distracted drivers to the risk of crashing and/or intervene with momentary braking or steering to avoid/mitigate crashes. Front crash prevention (FCP) and lane departure prevention (LDP) systems, in particular, address rear-end, sideswipe, single-vehicle, and lane drift crashes that are highly associated with distraction [6, 7]. FCP includes forward collision warning (FCW) and automatic emergency braking (AEB) features that effectively reduce front-to-rear crashes, as evident by a 27% reduction with FCW alone and a 50% reduction when combined with AEB. FCP technology has also demonstrated effectiveness in reducing crash rates of heavy trucks and crashes involving pedestrians or cyclists. Opportunities for further research also include examining the association between crash avoidance interventions and driving distractions, including the investigation of whether drivers use these features to engage in distracting activities.

### *2.4 Automated Enforcement Based Prevention*

Automated enforcement of traffic laws, including those proscribing certain distracting behaviors, is another technology that could be deployed against distracted driving. Automated enforcement deters drivers from engaging in targeted behaviors by helping drivers understand that they are likely to be sanctioned even when enforcement officers cannot possibly observe all offenses. For example, Australia began using cameras mounted over traffic lanes to detect violations of laws prohibiting handheld electronic device use [8]. Image recognition and machine-learning techniques process the images of vehicles with offending drivers in real time.

While, there is no firm link between distracted driving and red-light running or speeding in the record of real crashes, surveys and simulator studies suggest that inattention may lead to these behaviors. Research has shown significant reductions in red-light violations and speeding after introducing camera programs, leading to fewer crashes. While the effectiveness of automated enforcement in directly addressing distracted driving is still uncertain, there's potential for it to mitigate distracting behaviors like texting while driving.

### *2.5 Policy and Regulation*

Public policy has attempted to address driver distraction through different mechanisms and various authorities, from local municipalities to various branches of the federal government. The National Highway Traffic Safety Administration issued guidance concerning the electronic equipment in vehicles, and portable aftermarket devices brought into the vehicle. The Bipartisan Infrastructure Law requires (§24209) the Secretary of Transportation to research the use of DMS to mitigate driver distraction, and if deemed appropriate, to issue a rulemaking to require such systems. In Europe, DMS technologies are expected to become standard features in new cars as a result of regulatory and rating agency requirements. For example, the European Union has mandated drowsiness and attention monitoring and advanced driver distraction warnings for inclusion in all new vehicle models starting in 2024 and 2026 respectively. Additionally, the Euro NCAP grants vehicle points toward a 5-star rating for including DMS from 2023.

The automotive industry also developed consensus guidelines regarding designing and constructing in-vehicle electronic devices. In 2006, the Alliance of Automobile Manufacturers developed its Statement of Principles, Criteria, and Verification Procedures on Driver Interactions with Advanced In-Vehicle Information and Communication Systems. Updated in 2021, the guidelines, under the purview of the Alliance for Automotive Innovation, were amended to account for new technologies to assess driver attention and engagement. Outstanding research gaps include evaluating the cultural, societal, and regulatory differences across various countries and how these differences may impact the implementation and acceptance of technologies to mitigate distraction.

## 3. Discussion and Conclusions

This work summarizes existing approaches and technologies to prevent driver distraction and mitigate its consequences. It also identifies gaps in current technology and outlines areas for future research and development. The work covers technologies like DMS, smartphone-blocking technologies, crash avoidance features, and automated enforcement. While these technologies show promise, further research is needed to assess their effectiveness and adoption. From a regulatory standpoint, there are currently no specific requirements in the U.S. for technologies to mitigate driver distraction, but guidelines exist to minimize it. The Bipartisan Infrastructure Law mandates research into DMS technologies as a countermeasure to distraction. The work underscores the importance of understanding human-technology interaction and calls for continued research on automation and driver distraction. Innovative technological solutions have the potential to reduce driver distraction and improve road safety, but further development and understanding are necessary.

## 4. Acknowledgments

The authors extend gratitude to the National Distracted Driving Coalition for the organization and support. The views and conclusions expressed are those of the authors and have not been sponsored, approved, or endorsed by their affiliated organizations.

## References

[1] Gershon, P., Sita, K. R., Zhu, C., Ehsani, J. P., Klauer, S. G., Dingus, T. A., & Simons-Morton, B. G. (2019). Distracted driving, visual inattention, and crash risk among

teenage drivers. *American journal of preventive medicine*, 56(4), 494-500.

[2] National Center for Statistics and Analysis. (2022, May). Distracted Driving 2020 (Research Note. Report No. DOT HS 813 309). *National Highway Traffic Safety Administration*.

[3] Gershon, P., Zhu, C., Klauer, S. G., Dingus, T., & Simons-Morton, B. (2017). Teens' distracted driving behavior: Prevalence and predictors. *Journal of safety research*, 63, 157-161.

[4] Norman, D. A. (1990). The 'problem' with automation: inappropriate feedback and interaction, not 'over-automation'. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 327(1241), 585-593.

[5] Reagan, I. J., Teoh, E. R., Cicchino, J. B., Gershon, P., Reimer, B., Mehler, B., & Seppelt, B. (2021). Disengagement from driving when using automation during a 4-week field trial. *Transportation research part F: traffic psychology and behavior*, 82, 400-411.

[6] Cicchino, J. B., & Zuby, D. S. (2017). Prevalence of driver physical factors leading to unintentional lane departure crashes. *Traffic injury prevention*, 18(5), 481-487.

[7] Fildes, B., Keall, M., Bos, N., Lie, A., Page, Y., Pastor, C., ... & Tingvall, C. (2015). Effectiveness of low speed autonomous emergency braking in real-world rear-end crashes. *Accident Analysis & Prevention*, 81, 24-29.

[8]. Thompson, J. P., & Wundersitz, L. N. (2024). Mobile phone detection cameras in South Australia. *Prevention*, 119, 225-236.

# Asynchrony of Directional Auditory Warnings and Visual Information in Driving: Analysis of Eye Movement Patterns and Driver Responses

Ali Arabian [1*], Courtney M. Goodridge[1], Natasha Merat[1]

[1] Institute for Transport Studies, University of Leeds, Leeds, United Kingdom
tsaar@leeds.ac.uk

**Abstract:**
This study aims to investigate how directional auditory warnings affect drivers' gaze patterns and responses, and if this is affected by audiovisual asynchrony. A total of 44 drivers are recruited for this driving simulator study, which will assess how drivers respond to a range of warnings after resuming control from automation, during a collision avoidance task. Data collection is currently in progress. Insights gained from this experiment will improve our understanding of how directional auditory warnings guide drivers' visual attention, after resuming control from automation, particularly when these warnings are presented before the appearance of a hazard. Moreover, the findings will identify the optimal method for using directional warnings to guide attention in a lane change scenario; and whether these should be towards steering or hazard direction.

## 1. Introduction

The distribution of driver visual attention is an important aspect of road safety, with inattention towards the driving environment being implicated in a significant majority of traffic accidents. For instance, visual inattention is known to be a contributing factor in 93% of all rear-end collisions (Dingus et al., 2006) and 65% of all accidents (Graab et al., 2008). Visual attention is the cognitive process that facilitates the selection of important information from the environment. (Lockhofen & Mulert, 2021) and understanding the mechanisms underlying visual attention is vital for designing effective warning systems that can enhance driver safety. Shifts of visual attention may occur overtly through eye movements or covertly through cognitive focus, without eye movements (Posner, 1980). However, according to Moray (1993), visual information sampling is mostly influenced by overt attention. Attentional control mechanisms are classified broadly as bottom-up and top-down (Borji & Itti, 2012). While bottom-up mechanisms are involuntary and driven by the saliency of stimuli (Itti & Koch, 2000; Theeuwes, 1994), top-down mechanisms are voluntary and shaped by the observer's goals and knowledge (Chun & Jiang, 1998; Wolfe et al., 1989; Saalman et al., 2007).

Individuals employ different mechanisms to select important visual information from the environment, during real world tasks such as driving. However, they tend to select relevant information from a common external source across several modalities, thus their attention is crossmodally coordinated (Driver & Spence 1998). As an example, when we hear a car horn, we tend to look at the direction of the sound, even if that direction is not so visually salient. As a result, hearing a salient sound can draw visual-spatial attention towards its origin. A recent review in this area provide evidence that directing attention towards a peripheral, salient, sound improves visual perception, increases visual-cortical responses, and modulates visual cortex activity (Störmer, 2019). Notably, these results were contingent upon presenting the sound before the appearance of a visual stimulus. These findings have motivated the use of directional auditory warnings to guide drivers' visual attention. In the context of driving, directional warnings are used to refer to the hazard direction or the steering direction, particularly in a lane change scenario. The use of these directional auditory warnings is becoming more widespread through advanced driver assistance systems (ADAS) such those used for pedestrian collision warning (Chen et al., 2022), or blind spot warning (de Winter et al., 2022).

Several studies have investigated the effectiveness of auditory warnings that point to the hazard location (hazard direction), versus steering away from the hazard (steering direction), by identifying the direction of free space (Chen et al., 2022; Huang & Pitts, 2022; Wang et al., 2003). However, these studies have yielded inconsistent results. Some studies have reported that steering away from the hazards results in a faster steering response (Huang & Pitts, 2022; Wang et al., 2003), while Chen et al., 2022 found that directing attention to the hazard led to a faster response. Notably, none of these studies included presenting the directional auditory warning before the appearance of the visual stimulus, which was previously introduced as the main factor for the effectiveness of guiding visual attention (Störmer, 2019). De Winter et al. (2022) conducted a computer-based study to investigate how auditory hazard direction versus steering direction warnings guided drivers' gaze during lane change scenarios. Participants watched videos of near collisions on a computer screen and were asked to press the left or right arrow keys on the keyboard to indicate which direction they would move the vehicle, in order to avoid a collision. The presentation of auditory warnings coincided with the appearance of the hazard as the visual stimulus. Results revealed no significant differences in reaction times between steering and hazard direction warnings. Eye-tracking data indicated that directional auditory warnings did not guide drivers' visual attention towards the suggested location. The authors concluded that visual information is dominant and more influential for guiding attention, than auditory directional warnings. These results are predicted by the Colavita visual dominance effect (Colavita, 1974), which states that people frequently fail to respond to auditory stimuli if they have to respond to a simultaneously presented visual stimulus (Colavita, 1974; Spence et al., 2012). To modulate the Colavita visual dominance effect, Koppen and Spence (2007) conducted a study using audiovisual asynchrony. They manipulated the stimulus onset asynchrony (SOA) between the auditory and visual elements of bimodal targets. Participants were asked to quickly respond to auditory tones and visual flashes using separate keys. The findings indicated that the Colavita effect disappeared when participants received the auditory stimulus before the visual stimulus. To study this phenomenon in driving, the current study considers how the asynchrony of directional auditory warnings and visual information in driving affects eye movement patterns and driver responses, which to the best of our knowledge, has not been considered, to date.

## 2. Method

The study has received approval from the University of Leeds Ethics Committee (Reference code: 2024-1012-1510).

### *2.1 Participants*

A total of 44 participants aged between 21 and 70 years old will take part in this study. They will all hold a UK driving license, valid for at least 2 years. They will be regular drivers, who drive at least once a week. They will have normal or corrected to normal levels of vision and hearing.

### *2.2 Apparatus*

The study will be conducted using the University of Leeds static driving simulator (Figure 1). Eye tracking data will be collected via a Smart Eye Pro 3-camera fixed-based eye tracker. Participants will wear a headphones throughout the study, which is used to present the auditory warnings (see below for further details).

Figure 1: Static driving simulator set up.

### *2.3 Experimental design*

A 2 x 3 x 4 within-participants design will be utilised in the current study with audio warning type (speech and non-speech), warning direction (hazard direction, steering direction, and baseline/non-directional), and SOA (0, 200, 400, and 600 ms) as the independent variables. Each participant will complete four drives. The drives with speech and non-speech warnings will be counterbalanced across participants. Each drive will consist of 24 randomized trials regarding hazard location (left vs. right), directional and non-directional warnings, and SOAs. As an example, for the speech trials, drivers will hear the word "left" in their left ear, "right" in their right ear, or "look" for the non-directional warnings (see Tables 1 and 2).

### *2.4 Procedure and driving scenario*

After arrival, participants will sign and read a consent and information form. The form indicates the study purpose and explains the procedure and warnings. Prior to each main drive, participants will take part in a practice drive to become familiar with the equipment, scenarios, and auditory warnings. Practice drive trials will not include audiovisual asynchrony (SOA = 0). Each experimental drive begins with automated SAE Level 2 driving engaged (SAE, 2021). When the trial starts, the ego vehicle will travel at 60 mph on a three-lane highway, driving in the centre of the middle lane. After 4 seconds of automated driving, an occlusion is introduced, where the main driving scene and all mirrors are occluded (with a grey screen) for 4 s + the SOA. Before the end of each occlusion, one of the auditory warnings is presented via a set of headphones (e.g., left, right or look, for the speech trials). The onset of the auditory warning coincides with the onset of an SOA. For example, in a trial with an SOA of 0 ms, the appearance of the driving environment coincides with the onset of the auditory warning, and in a trial with an SOA of 600 ms, the driving environment appears 600 ms after the onset of the auditory warning. Participants are not aware of the SOAs. When the occlusion ends, the automation is turned off and a stationary vehicle appears ahead of the ego vehicle blocking the middle lane. Additionally, an offside truck will be travelling at a speed of 70 mph, in either the left or right lane. The presence of this truck precludes drivers from moving to the adjacent lane (the lane with the truck), forcing them to move to the free lane, to avoid hitting the stationary lead vehicle. Participants must avoid a collision by steering to the free lane within a time to collision (TTC) of 3 seconds. The trial then ends, and participants are asked to press the button on the steering wheel to initiate the next trial. Regarding the warnings, two categories are used: *steering direction* warnings require participants to steer towards the free lane, and *hazard direction* warnings inform participants about the lane with the truck. At the end of each block, participants will fill out a post-drive questionnaire assessing their acceptance of warnings, perceived urgency of warnings, and NASA-TLX scores.

## 3. Results

Data collection is currently in progress. If accepted, the presentation will delve into the variations in gaze distribution, steering response time, collision rate, maximum lateral acceleration, and time to lane change in response to the different audiovisual asynchronies and whether these are different for the different warning directions and types. Regression-based models will be used to investigate differences in responses between the varying conditions.

## 4. Discussion and Conclusions

The study's results will be discussed, and the potential implications of the asynchrony between the onset of auditory warnings and the appearance of visual information on drivers' gaze patterns and responses in a critical lane change scenario will be outlined. Insights gained from this study hold significant implications for the design and implementation of future ADAS, aiming to enhance road safety and mitigate traffic accidents, especially regarding obstacle avoidance after resumption of control in Level 2 automated driving.

## 5. Acknowledgments

The authors would like to thank the Simulator Team of the Institute for Transport Studies, University of Leeds, especially Albert Solernou Crusat, and Michael Daly for creating the simulated environments. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101006664.

***Table 1*** *Number of trials per drive and the number of drives per speech block.*

| Audio direction | Steering direction drive (24 trials) | | | | Hazard direction drive (24 trials) | | | |
|---|---|---|---|---|---|---|---|---|
| | "Left" from left headphones | "Right" from right headphones | "Look" from both headphones (Truck in right) | "Look" from both headphones (Truck in left) | "Left" from left headphones | "Right" from right headphones | "Look" from both headphones (Truck in right) | "Look" from both headphones (Truck in left) |
| Driver response | Steer left (8 trials) | Steer right (8 trials) | Steer left (4 trials) | Steer right (4 trials) | Steer right (8 trials) | Steer left (8 trials) | Steer left (4 trials) | Steer right (4 trials) |
| SOAs | | | | | | | | |
| 0 (6 trials) | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |
| 200 (6 trials) | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |
| 400 (6 trials) | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |
| 600 (6 trials) | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |

***Table 2*** *Number of trials per drive and the number of drives per non-speech block.*

| Audio direction | Steering direction drive (24 trials) | | | | Hazard direction drive (24 trials) | | | |
|---|---|---|---|---|---|---|---|---|
| | "Beep" from left headphones | "Beep" from right headphones | "Beep" from both headphones (Truck in right) | "Beep" from both headphones (Truck in left) | "Beep" from left headphones | "Beep" from right headphones | "Beep" from both headphones (Truck in right) | "Beep" from both headphones (Truck in left) |
| Driver response | Steer left (8 trials) | Steer right (8 trials) | Steer left (4 trials) | Steer right (4 trials) | Steer right (8 trials) | Steer left (8 trials) | Steer left (4 trials) | Steer right (4 trials) |
| SOAs | | | | | | | | |
| 0 (6 trials) | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |
| 200 (6 trials) | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |
| 400 (6 trials) | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |
| 600 (6 trials) | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |

## References

Borji, A., & Itti, L. (2012). State-of-the-art in visual attention modeling. IEEE transactions on pattern analysis and machine intelligence, 35(1), 185-207. https://doi.org/10.1109/TPAMI.2012.89

Chen, J., Šabić, E., Mishler, S., Parker, C., & Yamaguchi, M. (2022). Effectiveness of lateral auditory collision warnings: should warnings be toward danger or toward safety? Human factors, 64(2), 418-435. https://doi.org/10.1177/0018720820941618

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. Cognitive psychology, 36(1), 28-71. https://doi.org/10.1006/cogp.1998.0681

Colavita, F. B. (1974). Human sensory dominance. Perception & psychophysics, 16(2), 409-412. https://doi.org/10.3758/BF03203962

de Winter, J., Hu, J., & Petermeijer, B. (2022). Ipsilateral and contralateral warnings: effects on decision-making and eye movements in near-collision scenarios. Journal on Multimodal User Interfaces, 16(3), 303-317. https://doi.org/10.1007/s12193-022-00390-6

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. (2006). The 100-car naturalistic driving study, Phase II-results of the 100-car field experiment. https://rosap.ntl.bts.gov/view/dot/37370/dot_37370_DS1.pdf

Driver, J., & Spence, C. (1998). Attention and the crossmodal construction of space. Trends in cognitive sciences, 2(7), 254-262. https://doi.org/10.1016/S1364-6613(98)01188-7

Graab, B., Donner, E., Chiellino, U., & Hoppe, M. (2008). Analyse von Verkehrsunfällen hinsichtlich unterschiedlicher Fahrerpopulationen und daraus ableitbarer

Ergebnisse für die Entwicklung adaptiver Fahrerassistenzsysteme. 3. Tagung Aktive Sicherheit durch Fahrerassistenz, https://mediatum.ub.tum.de/doc/1145118/document.pdf

Huang, G., & Pitts, B. J. (2022). Takeover requests for automated driving: The effects of signal direction, lead time, and modality on takeover performance. Accident analysis & prevention, 165, 106534. https://doi.org/10.1016/j.aap.2021.106534

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. Vision research, 40(10-12), 1489-1506. https://doi.org/10.1016/S0042-6989(99)00163-7

Koppen, C., & Spence, C. (2007). Audiovisual asynchrony modulates the Colavita visual dominance effect. Brain research, 1186, 224-232. https://doi.org/10.1016/j.brainres.2007.09.076

Lockhofen, D. E. L., & Mulert, C. (2021). Neurochemistry of visual attention. Frontiers in neuroscience, 15, 643597. https://doi.org/10.3389/fnins.2021.643597

Moray, N. (1993). Designing for attention. https://psycnet.apa.org/record/1994-97378-006

Posner, M. I. (1980). Orienting of attention. Quarterly journal of experimental psychology, 32(1), 3-25. https://doi.org/10.1080/00335558008248231

Saalmann, Y. B., Pigarev, I. N., & Vidyasagar, T. R. (2007). Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. Science, 316(5831), 1612-1615. https://doi.org/10.1126/science.1139140

SAE. (2021). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. SAE international, 4970(724), 1-5. https://www.sae.org/standards/content/j3016_202104/

Spence, C., Parise, C., & Chen, Y.-C. (2012). The Colavita visual dominance effect. The neural bases of multisensory processes. https://www.ncbi.nlm.nih.gov/books/NBK92851/?report=reader

Störmer, V. S. (2019). Orienting spatial attention to sounds enhances visual processing. Current opinion in psychology, 29, 193-198. https://doi.org/10.1016/j.copsyc.2019.03.010

Theeuwes, J. (1994). Endogenous and exogenous control of visual selection. Perception, 23(4), 429-440. https://doi.org/10.1068/p230429

Wang, D.-Y. D., Proctor, R. W., & Pick, D. F. (2003). Stimulus-Response Compatibilitiy Effects for Warning Signals and Steering Responses. Driving Assessment Conference, https://doi.org/10.17077/drivingassessment.1128

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: an alternative to the feature integration model for visual search. Journal of Experimental Psychology: Human Perception and Performance, 15(3), 419. https://psycnet.apa.org/doi/10.1037/0096-1523.15.3.419

# Driver Monitoring Based Lane Keeping Aid for Improved User Acceptance

Daban Rizgary [1*], Niklas Strand [1], Victor Brandt [2], Erik Larsson [3], Ebba Davidsson [3], Fabian Burman [3]

[1] RISE Research Institutes of Sweden, Division Digital Systems, Mobility and Systems, Gothenburg, Sweden
[2] Smart Eye, Gothenburg, Sweden
[3] Aptiv, Mölndal, Sweden
[*]daban.rizgary@ri.se

**Abstract: This research examines the utility of integrating Driver Monitoring Systems (DMS) and ambient light displays into Lane Keeping Aids (LKAs) to reduce the incidence of false warnings. Given the substantial role driver inattention plays in road accidents, with distractions from Non-Driver-Related-Tasks (NDRTs), the study explores how customizing LKA functions based on the driver's attention can improve its user acceptance by reducing false positive warnings. An experimental setup involving 8 expert participants was conducted on a 26-kilometer stretch of public road to assess the impact of four different LKA configurations: standard LKA, LKA supplemented with ambient light, Adaptive LKA that takes driver's attention into account (A-LKA), and A-LKA together with ambient light. The findings reveal that the adaptive LKA systems, particularly when used in combination with ambient light, were preferred for their ability to minimize false positives, while improving effectiveness of true positive warnings. The study underscores the potential of adapting LKA systems to individual drivers' attention levels and the role of ambient light in such systems.**

## 1. Introduction

### *1.1 DMS-based lane keeping aid*

In the context of road safety research, the issue of driver distraction and inattention has emerged as a critical factor contributing to vehicular accidents. Accident data from 6 European countries has shown that in 32% out of 1005 crash incidents, the driver, a rider or a pedestrian was either distracted or inattentive (Talbot, Fagerlind & Morris, 2013).

The work Pohl et al. (2007) and Blaschke et al. (2009) investigate the role of driver monitoring systems (DMS) in reducing false positive warnings in lane keeping systems. These two studies explore the effect of drivers' Non-Driver-Related-Tasks (NDRTs) on lane keeping and show that most NDRTs increase lane deviations, but that in most cases drivers still manage to stay in the lane.

The studies show that all degrees of lateral support increase lane keeping performance, with systems providing greater assistance during NDRT tasks proving most useful. This underlines the potential of customizing lane keeping aid (LKA) systems based on the driver's use of NDRT, which could contribute to the reduction of false positive warnings i.e warnings that occur when driver is aware and alert, and thus improve driver safety and comfort (Pohl et al., 2007) and (Blaschke et al., 2009).

A driver monitoring system (DMS) that can adapt to the driver's attention level and behaviour could reduce the need for warnings from lane-keeping systems and thus reduce the number of false positives. This in turn might lead to less deactivation of the feature by the driver and a more positive attitude towards it. By integrating such systems, lane-keeping aid could be provided in a more targeted and efficient way, increasing both road safety and driver comfort.

### *1.2 Ambient light displays*

Advances in cognitive psychology have led to various theories explaining how humans process information, especially when it comes to performing multiple tasks simultaneously. One of the debates concerns the single-channel versus multiple resources hypothesis (Navon & Miller, 2002; Wickens, 2002; Wickens, 2008). The theory of multiple resources, proposed by Wickens (Wickens, 2002), suggests that there are two aspects of visual processing, i.e. focal and peripheral vision. Moreover, these two aspects are suggested to support efficient time-sharing and are characterized by different neurological structures (Previc, 1998). (Borojeni et al., 2016) state that this theory highlights a potential bottleneck in visual processing tasks during take-over requests (TORs) in automated driving, suggesting the usefulness of peripheral light displays to facilitate efficient time-sharing for TORs.

The concept of ambient light displays in a vehicle context is mainly about using non-intrusive, peripheral light signals to convey information to drivers. These signals can offer guidance to the driver in tasks such as lane changes and TOR in semi-autonomous vehicles. Studies like (Löcken et al., 2015) and (Bojoreni et al., 2016) delve into the application of ambient light in specific driving situations, and have provided insights into how ambient light signals can improve driver performance without excess cognitive load for the driver.

In the case of lane departure warnings, which often use icons, sounds and tactile information, the visual aspect can be emphasized by using ambient light displays instead of icons. Ambient light displays can be placed on each A-pillar relating to the respective lane markings (see figure 1), thus potentially having the warnings be more comprehensible.

## 2. Method

An experiment was conducted on a 26-kilometre stretch of public road using a Lincoln MKZ test vehicle and a proof-of-concept prototype for LKA and an ambient light Human-Machine Interface (HMI).

The study was conducted with four different conditions: standard LKA which showed lane departure warnings with an icon in the dashboard, standard LKA with ambient light, A-LKA (Adaptive LKA) in which drivers' attentiveness was monitored by means of a camera-based DMS and attentiveness was considered for LKA warning and intervention triggers. When driver distraction was detected, the LKA was more prone to produce warnings and intervention. A-LKA showed lane departure warnings with an icon in the dashboard. The fourth and last condition was A-LKA with ambient light. Each driver, a total of 8 males with an average age of 35 years (SD = 8.3) who were test drivers at research project partner Aptiv, completed a 26 kilometres drive under each condition on the same public roadway that included about half the stretch rural roads and the other half highway, resulting in a total of four 26-km drives per participant. On the highway, the participants were instructed to perform a manual distraction task by placing one hand around the infotainment area of the vehicle with eyes on the road whilst slowly drifting towards the lane line to experience a LKA warning and intervention. This technique was used to avoid danger during the data collection while still providing the driver with an experience of LKA triggers with and without the A-LKA functionality in the highway context. On the rural road, the participants were instructed to drive naturally. Given the nature of the rural road which was selected for its narrower road width, lane departure warnings were triggered in all drives.

To evaluate the systems, several subjective rating measurement were used, including the Van der Laan Acceptance Scale containing usefulness and satisfaction subscales, individual rating questions on effectiveness (How clear was the driver assistance system in its communication to you about what it saw and did? ), perceived situation awareness (The driver assistance system helped me understand my surroundings better), trust (How much do you trust the system?), preference (I am convinced that driver assistance systems should work this way) and ease of use (Generally, this/these features were (1- very difficult to use and 7 - very easy to use), as well as workload as measured by a revised NASA-RTLX scale. Semi-structured interviews were also conducted after each condition with questions about the LKA logic and behavior, as well as the user interface information and design.

The study focused on comparing the different conditions to investigate how the addition of ambient light and the adaptation of driver support functionality to the driver's attention level affected drivers in the mentioned measures.

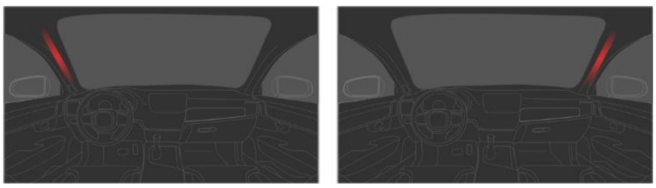


*Figure 1: illustration of the location and shape of ambient light that showed the approach or breach of the lane line or lane edge.*

## 3. Results

The interviews conducted showed that LKA with ambient light was reported as more effective in its communication than LKA alone, but also that it led to an increased awareness of false warnings. Drivers were very positive about the idea of adapting LKA to the driver's condition and preferred the balance achieved by minimizing false positives and clearly presenting true positives with ambient light (A-LKA + ambient).

Inferential analysis of the subjective ratings showed that:

- Usefulness: A-LKA with ambient light was rated as significantly more useful than LKA ($F(3, 21) = 4.480$, $p = 0.014$), with a large effect size (Cohen's d = -1.723).

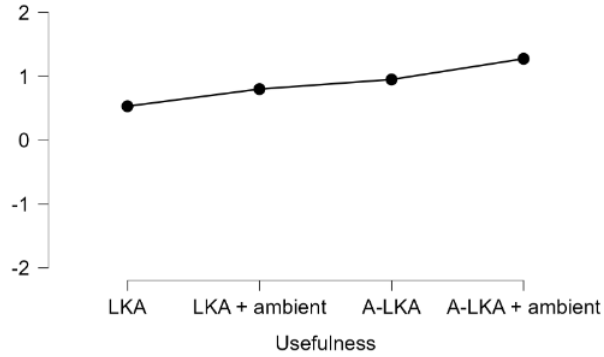



*Figure 2: line graph showing mean score difference between different conditions for usefulness.*

- Satisfaction: Both A-LKA and A-LKA with ambient light were significantly more satisfactory than LKA with ambient light ($F(3, 21) = 7.103$, $p = 0.002$), with large effect sizes (Cohen's d = -1.718).

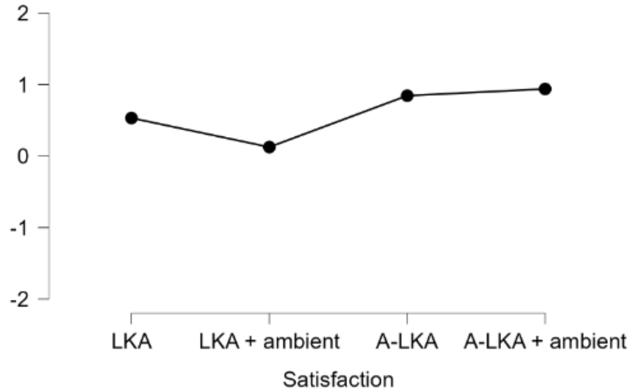



*Figure 3: line graph showing mean score difference between different conditions for satisfaction.*

- Effectiveness: LKA with ambient light outperformed both A-LKA and A-LKA with ambient light ($F(3, 21) = 38.979$, $p < .001$), with large effect sizes (Cohen's d = -2.269). A-LKA with ambient light was more effective than A-LKA alone (Cohen's d = -2.695).

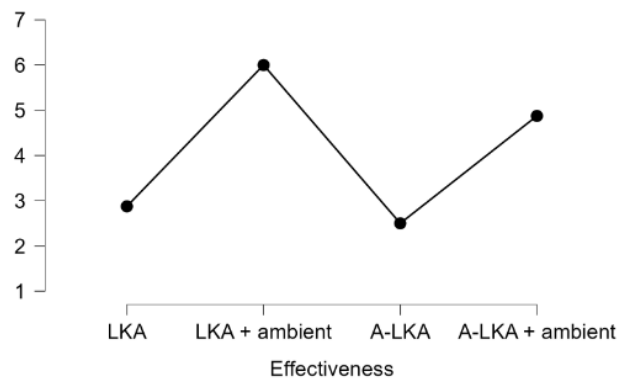



*Figure 4: line graph showing mean score difference between different conditions for effectiveness.*

- Perceived situation awareness: A-LKA with ambient light showed a tendency towards higher perceived situation awareness compared to LKA and A-LKA alone ($F(3, 21) = 3.098$, $p = 0.049$), though not statistically significant (Cohen's d = -0.854).

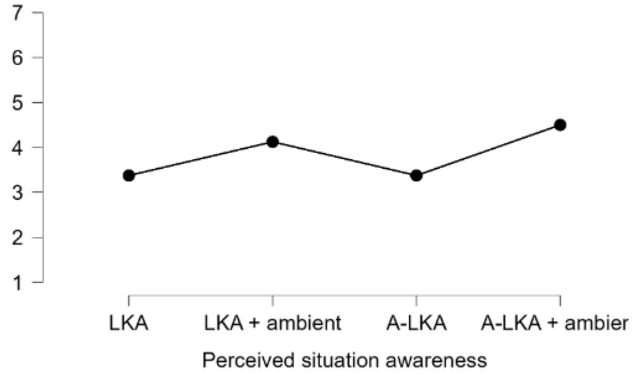



*Figure 5: line graph showing mean score difference between different conditions for perceived situation awareness.*

- Preference: Participants significantly preferred A-LKA with ambient light over LKA ($F(3, 21) = 3.363$, $p = 0.038$), with a large effect size (Cohen's d = -1.452).

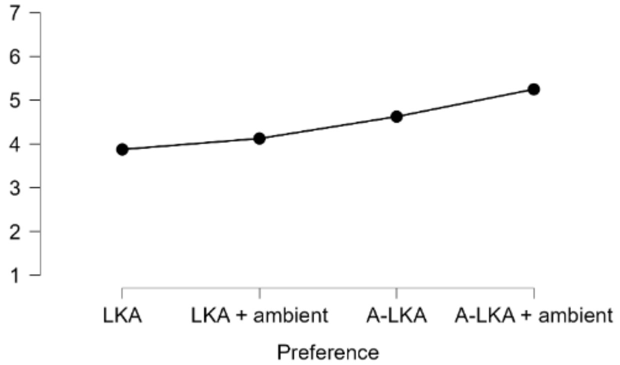



*Figure 6: line graph showing mean score difference between different conditions for preference.*

- No differences were observed in ease of use ($F(3, 21) = 1.449$, $p = 0.257$), trust ($F(3, 21) = 2.116$, $p = 0.129$), or workload (NASA-RTLX: $F(3, 21) = 1.420$, $p = 0.265$).

  See Appendix (A) for numeric descriptive results.

## 4. Conclusions

The adaptive LKA, especially in combination with ambient light, seems to offer several advantages over the basic LKA. However, even the basic LKA with ambient light is still assessed as clear (in its communication), emphasizing that ambient light clarifies communication in the LKA systems. The adaptive nature of the warnings, which are adjusted based on the driver's attention, seems to be appreciated by drivers in terms of usability, satisfaction, and overall preference.

This study demonstrates the feasibility of incorporating driver monitoring data for avoiding excessive LKA triggers in situations where the driver is attentive, alert and aware of the driving situation. This study also demonstrates the potential utility of ambient light for LKA warnings as a means to increase the effectiveness of the warning. These types of solutions have the potential to increase user acceptance of LKA, but in order to avoid missing true positives, the DMS needs to account for situations where the driver is looking but not seeing e.g. during a state of mind wandering.

## 5. Acknowledgments

We would like to thank Sweden's FFI - Strategic Vehicle Research and Innovation program for funding this research.

## 6. References

1. Blaschke, C., Breyer, F., Färber, B., Freyer, J., & Limbacher, R. (2009). Driver distraction based lane-keeping assistance. *Transportation research part F: traffic psychology and behaviour*, 12(4), 288-299.
2. Borojeni, S. S., Chuang, L., Heuten, W., & Boll, S. (2016). Assisting drivers with ambient take-over requests in highly automated driving. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.* https://doi.org/10.1145/3003715.3005409
3. Locken, A., Muller, H., Heuten, W., & Boll, S. (2015). An experiment on ambient light patterns to support lane change decisions. *2015 IEEE Intelligent Vehicles Symposium (IV)*. https://doi.org/10.1109/ivs.2015.7225735
4. Navon, D., & Miller, J. (2002). Queuing or sharing? A critical evaluation of the single-bottleneck notion. *Cognitive Psychology*, *44*(3), 193–251. https://doi.org/10.1006/cogp.2001.0767
5. Pohl, J., Birk, W., & Westervall, L. (2007). A driver-distraction-based lane-keeping assistance system. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 221(4), 541-552.
6. Previc, F. H. (1998). The neuropsychology of 3-D space. *Psychological Bulletin*, 124(2), 123–164. https://doi.org/10.1037/0033-2909.124.2.123
7. Talbot, R., Fagerlind, H., & Morris, A. (2013). Exploring inattention and distraction in the Safetynet Accident Causation Database. *Accident Analysis & Prevention*, 60, 445–455. https://doi.org/10.1016/j.aap.2012.03.031
8. Wickens, C. D. (2002). Multiple Resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806
9. Wickens, C. D. (2008). Multiple Resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 449–455. https://doi.org/10.1518/001872008x288394

## Appendix A

**Table 1** Descriptive data for subjective rating scores separated by measures and conditions. Van Der Laan (usefulness and satisfaction) were scored on a scale -2 to 2, the individual rating questions (effectiveness, perceived situation awareness, trust, preference and ease of use) were scored on a 1 to 7 scale, and the modified NASA-RTLX workload scale was scored on a 1 to 10 scale.

| Measure | | LKA | LKA + ambient | A-LKA | A-LKA + ambient |
|---|---|---|---|---|---|
| Usefulness | Median | 0.600 | 0.900 | 1.000 | 1.200 |
| Usefulness | Mean | 0.531 | 0.800 | 0.950 | 1.275 |
| Usefulness | Std. Deviation | 0.363 | 0.595 | 0.366 | 0.354 |
| Satisfaction | Median | 0.625 | 0.000 | 0.875 | 1.000 |
| Satisfaction | Mean | 0.531 | 0.125 | 0.844 | 0.938 |
| Satisfaction | Std. Deviation | 0.364 | 0.582 | 0.499 | 0.417 |
| Effectiveness | Median | 3.000 | 6.000 | 2.000 | 5.000 |
| Effectiveness | Mean | 2.875 | 6.000 | 2.500 | 4.875 |
| Effectiveness | Std. Deviation | 0.835 | 0.756 | 1.069 | 0.835 |
| Perceived situation awareness | Median | 3.000 | 4.000 | 3.000 | 4.500 |
| Perceived situation awareness | Mean | 3.375 | 4.125 | 3.375 | 4.500 |
| Perceived situation awareness | Std. Deviation | 0.916 | 1.246 | 1.506 | 1.512 |
| Trust | Median | 4.000 | 4.000 | 5.000 | 5.000 |
| Trust | Mean | 4.500 | 4.375 | 4.875 | 5.125 |
| Trust | Std. Deviation | 1.069 | 0.916 | 0.991 | 1.246 |
| Preference | Median | 4.000 | 4.000 | 4.500 | 5.000 |
| Preference | Mean | 3.875 | 4.125 | 4.625 | 5.250 |
| Preference | Std. Deviation | 0.991 | 1.126 | 0.744 | 0.886 |
| Ease of use | Median | 6.000 | 6.000 | 6.000 | 6.500 |
| Ease of use | Mean | 5.875 | 6.000 | 6.000 | 6.500 |
| Ease of use | Std. Deviation | 0.835 | 0.756 | 0.926 | 0.535 |
| NASA-RTLX | Median | 3.417 | 3.667 | 3.583 | 3.167 |
| NASA-RTLX | Mean | 3.417 | 3.750 | 3.646 | 3.333 |
| NASA-RTLX | Std. Deviation | 0.321 | 0.496 | 0.393 | 0.807 |

# Understanding individual differences in spare visual capacity and uncertainty growth: Insights from driving simulator studies

Hilkka Grahn [1*], Toni Hautaoja [1], Tuomo Kujala [1]

[1] Faculty of Information Technology, University of Jyväskylä, P.O. Box 3, FI-40014 University of Jyväskylä, Finland
[*]hilkka.grahn@jyu.fi

**Abstract: In the realm of driving, individuals possess spare visual capacity that is often directed towards engaging in secondary tasks. Nevertheless, considerable variability exists in individuals spare visual capacity. One method to estimate spare visual capacity involves the application of the visual occlusion technique, where the driver's visual field is intermittently blocked, and the occluded duration is measured. This occlusion time not only estimates spare visual capacity but measures the drivers' uncertainty growth rate. In this study, derived from a 4-year project, we present and summarize the findings from three driving simulator studies, providing a unique perspective on their results. Here, our objective is to investigate how drivers' individual differences impact uncertainty growth rate and to identify relevant contributing factors. Results indicate that drivers base their occlusion time choices on a personal preference level, potentially linked to individual uncertainty tolerance or uncertainty growth rate during occlusion, observed in both manual and assisted driving. Also, it seems that drivers' subjective estimates of their spare visual capacity display more variability than more objective estimates of their spare visual capacity. Furthermore, various factors seem to contribute to occlusion time preferences, with risk-taking tendencies being a noteworthy area requiring further investigation. Despite these findings, the most significant individual factors influencing occlusion times and spare visual capacity in driving remain unknown. These results can be utilized, for example, in the development of individualized and contextually intelligent driver attention monitoring and warning systems.**

## 1. Introduction

Individuals possess spare visual capacity while driving, as indicated by Safford (1971). Spare visual capacity, as defined by Shinar et al. (1978, p. 556), is "the inverse of the amount of time a driver must maintain his eyes open in order to perform the driving task successfully". Spare visual capacity is frequently allocated to engaging in secondary tasks (Kircher & Ahlström, 2018), and Hoedemaeker and Kopf (2001) and Mars et al. (2014) suggest that it can be enlarged through driving automation. One method to estimate spare visual capacity involves employing the visual occlusion technique, wherein driver's visual field is intermittently blocked (Kujala et al., 2023; Senders et al., 1967). However, there are differences among individuals in terms of the extent of their spare visual capacity (Safford, 1971). Similarly, it has been found that individual differences also influence drivers' occlusion times (Kujala, Grahn, et al., 2016; Kujala, Mäkelä, et al., 2016).

Furthermore, occlusion can serve as a tool to measure drivers' uncertainty growth rate (Senders et al., 1967). While secondary tasks have been employed for a similar purpose, we propose that occlusion offers a more direct assessment of drivers' uncertainty adaptation. This distinction arises because certain features of secondary tasks, such as their structure (Grahn & Kujala, 2020), can cause prolonged in-car glances. Here, our objective is to investigate how drivers' individual differences impact uncertainty growth and identify relevant factors. The first study utilizes both a visual secondary task and occlusion, while the subsequent two studies rely solely on occlusion.

## 2. Empirical foundation

The empirical foundation is derived from a 4-year project, "Appropriate Uncertainty in Manual and Automated Driving", focused on creating computational models to improve traffic safety. These models define drivers' situational information sampling requirements, contribute to the understanding of individual driver variability, and support the development of, for instance, valid and reliable inattention monitoring algorithms.

In this section, we explore three project-related studies (one published, one under review, and one in preparation), systematically analyzing each to offer summarized insights into the central theme of individual differences among drives. This summary provides a unique perspective on the combined findings of the three studies. These articles feature data from driving simulator studies, utilizing multilevel modeling for analysis, and only relevant results are reported here. Multilevel models, accommodating hierarchical data structures (Hox, 1998), account for individual variation through intraclass correlation (ICC) when a participant is included as a random effect. While additional variables were considered and controlled in each model, not all of them are detailed in this discussion.

### *2.1 Study 1*

In Study 1 (Grahn et al., 2023, see Figure 1), we investigated the association between drivers' occlusion times (OT) and in-car glance durations, while identifying factors influencing this association. With 30 participants in a driving simulator, the experiment comprised three driving tasks: self-paced occlusion and visually and cognitively demanding task (high and low demand), with the primary goal of lane-keeping.

Additionally, drivers' visual search speed was assessed while not driving. The self-paced occlusion drive estimated participants' preferred maximum OTs as their subjective estimates of spare visual capacity.

The findings indicate an association between individuals' preferences for OTs and in-car glance durations, particularly in low-demanding, unstructured tasks. Notably, in one model, longer visual search task duration in the stationary condition predicted longer glances during visual search tasks. The ICCs, reflecting the share of variance in the dependent variable attributed to individual differences, ranged from 15.9% to 29.7%.

*Figure 1: The setup of the experiment illustrating a visually and cognitively high-demanding secondary task (Grahn et al., 2023)*

### *2.2 Study 2*

In this study (Grahn et al., 2024a, see Figure 2), our aim was to investigate how drivers ($N = 30$) adapt their visual sampling behavior based on situational and individual variables. Participants engaged in a self-paced occlusion drive similar to Study 1. Alongside studying occlusion times (OT), we explored time-to-line-crossings (TLC, Godthelp et al., 1984), where OT reflects a driver's subjective estimate of spare visual capacity and TLC represents more objectively estimated spare visual capacity in a situation. Again, their primary task was to stay in their lane, while maximizing OT.

The findings indicate that the OT of the previous occlusion was the most influential predictor of current OT. Additionally, another significant predictor in the TLC models was standard deviation of lane position (SDLP), presenting individual variability in lateral control performance. Also, the ICCs in the models predicting OTs ranged from 24.7% to 59.5%. Notably, these ICCs were six to eleven times larger than those observed in models predicting TLC.

*Figure 2: The setup of the experiment: occlusion drive's unoccluded period (Grahn et al., 2024a)*

### *2.3 Study 3*

In Study 3 (Grahn et al., 2024b, see Figure 3), we investigated whether individual preferences for OT levels extend to assisted driving and also explored the potential association between OTs and general risk-taking behavior (Meertens & Lion, 2008). While intermittently driving occluded, 32 participants monitored an imperfect lane-keeping assist, contributing insights into how drivers adapt to and calibrate their uncertainty toward such systems based on system's observed behavior. Their task was to maximize OT while staying on the road.

Once again, the most influential factor was the OT of the previous occlusion. Also, the general tendency to take risks impacted OTs, with a higher inclination toward risk-taking being associated with longer OTs (on average 221 ms increase by 1-unit increase in the risk-taking scale) but was not significant in the model ($p = .076$). The ICC of the intercept-only model was 55.4% and in the final model 47.9%.

*Figure 3: The setup of the experiment: occlusion drive's unoccluded period (Grahn et al. 2024b)*

## 3. Discussion

### *3.1 Occlusion time preferences*

OT preferences were consistent across all three studies. In Study 1, there was an association between individuals' OT preferences and in-car glance durations, particularly in low-demanding, unstructured tasks. This suggests a consistent pattern of individual sampling preferences or tendencies during both occlusion drive and when engaging in tasks with minimal visual and cognitive demands. Study 2 found the OT of the previous occlusion as the most influential predictor of the current OT. This suggests that individuals base their OT choices on a personal preference level, potentially linked to individual uncertainty tolerance or uncertainty growth rate during occlusion. The same finding was observed in Study 3, indicating that even in assisted driving, individual preferences for OT choices remain valid.

### 3.2 Intraclass correlations across models

The ICCs of the models across studies were relatively high. The ICC measures the extent to which observations within the same individual are more similar to each other than to observations in other individuals. This strengthens the observation of individual OT preferences. Notably, in Study 2, the ICC in OT models was six to eleven times higher than in TLC models, indicating that subjective estimates of spare visual capacity vary much more than the more objective estimates of spare visual capacity in the same task.

### 3.3 Other individual factors

These studies also revealed insights into other individual factors influencing in-car glance durations, OT preferences, or TLC. Specifically, slower visual search speed was associated with longer in-car glances, while increased SDLP (instability in lateral control) increased the probability and rate of lane departure. Additionally, one of the models implied a potential association between a general risk-taking in life and longer OTs, although its statistical significance was slightly above the threshold in our study.

## 4. Conclusions

Individuals base their OT choices on a personal preference level, possibly linked to individual uncertainty tolerance or the uncertainty growth rate during occlusion. This was observed in both manual and assisted driving. Drivers' subjective estimates of their spare visual capacity (here: OT) exhibit more variability than the more objective estimates of their spare visual capacity (here: TLC). Furthermore, various factors seem to contribute to OT preferences. For example, the risk-taking tendencies merit further investigation. Overall, the most significant individual factors that influence drivers' individual OT preferences remain unknown.

## 5. Acknowledgments

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

## References

Godthelp, H., Milgram, P., & Blaauw, G. J. (1984). The development of a time-related measure to describe driving strategy. *Human Factors*, *26*(3), 257–268. https://doi.org/10.1177/001872088402600302

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., Hautaoja, T., & Salvucci D. (2024a). Investigating the situational dynamics of visual information sampling in lateral vehicle control – Subjective vs. objective estimates of spare visual capacity. [Manuscript submitted for publication].

Grahn, H., Kujala, T., Hautaoja, T., Srinivasan A., & Markkula G. (2024b). Drivers' adaptation and uncertainty calibration to imperfect lane-keeping assist: A driving simulator study. [Manuscript in preparation].

Grahn, H., Kujala, T., Taipalus, T., Lee, J., & Lee, J. D. (2023). On the relationship between occlusion times and in-car glance durations in simulated driving. *Accident Analysis and Prevention*, *182*, 106955. https://doi.org/10.1016/j.aap.2023.106955

Hoedemaeker, M., & Kopf, M. (2001). Visual sampling behaviour when driving with adaptive cruise control. *In Proceedings of the Ninth International Conference on Vision in Vehicles.*, 19–22.

Hox, J. (1998). Multilevel modeling: When and why. In I. Balderjahn, R. Mathar, & M. Schader (Eds.), *Classification, Data Analysis, and Data Highways. Studies in Classification, Data Analysis, and Knowledge Organization.* (pp. 147–154). Springer. https://doi.org/https://doi.org/10.1007/978-3-642-72087-1_17

Kircher, K., & Ahlström, C. (2018). Evaluation of methods for the assessment of attention while driving. *Accident Analysis and Prevention*, *114*, 40–47. https://doi.org/10.1016/j.aap.2017.03.013

Kujala, T., Grahn, H., Mäkelä, J., & Lasch, A. (2016). On the visual distraction effects of audio-visual route guidance. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 169–176. https://doi.org/10.1145/3003715.3005421

Kujala, T., Kircher, K., & Ahlström, C. (2023). A review of occlusion as a tool to assess attentional demand in driving. *Human Factors*, *65*(5), 792–808. https://doi.org/10.1177/00187208211010953

Kujala, T., Mäkelä, J., Kotilainen, I., & Tokkonen, T. (2016). The attentional demand of automobile driving revisited: Occlusion distance as a function of task-relevant event density in realistic driving scenarios. *Human Factors*, *58*(1), 163–180. https://doi.org/10.1177/0018720815595901

Mars, F., Deroo, M., & Hoc, J. M. (2014). Analysis of human-machine cooperation when driving with different degrees of haptic shared control. *IEEE Transactions on Haptics*, *7*(3), 324–333. https://doi.org/10.1109/TOH.2013.2295095

Meertens, R. M., & Lion, R. (2008). Measuring an individual's tendency to take risks: The risk propensity scale. *Journal of Applied Social Psychology*, *38*(6), 1506–1520. https://doi.org/10.1111/j.1559-1816.2008.00357.x

Safford, R. Jr. (1971). *Visual spare capacity in automobile driving and its sensitivity to carboxyhemoglobin* [Doctoral dissertation, The Ohio State University]. https://www.proquest.com/docview/302504873?pq-origsite=gscholar&fromopenview=true

Senders, J. W., Kirstofferson, A. B., Levison, W. H., Dietrich, C. W., & Ward, J. L. (1967). The attentional demand of automobile driving. *Highway Research Record*.

Shinar, D., McDowell, E. D., Rackoff, N. J., & Rockwell, T. H. (1978). Field dependence and driver visual search behavior. *Human Factors*, *20*(5), 553–559.

# Understanding hands-off patterns while using Tesla Autopilot: Individual differences and their impact on hands-off duration

Joonbum Lee [1*], Pnina Gershon [2], Josh Domeyer [3], Bruce Mehler [2], and Bryan Reimer [2]

[1] Department of Industrial and Systems Engineering, University of Wisconsin-Madison, 1415 Engineering Drive, Madison, USA
[2] MIT, Center for Transportation & Logistics, AgeLab, 1 Amherst Street, Cambridge, MA
[3] Toyota Collaborative Safety Research Center, 1555 Woodridge Ave., Ann Arbor, MI, USA
[*]joonbum.lee@wisc.edu

**Abstract: This study investigates the relationship between drivers' hand positions and visual attention while using partial driving automation to understand individual differences in hands-off behavior. We analyzed 298 instances of Tesla Autopilot disengagements. Results showed that longer hands-off durations were associated with reduced on-road glance times. Based on the relationship, two driver clusters were identified: Cluster 1, characterized by more frequent Autopilot use, engaged in a combination of hands-off and hands-on behavior, while Cluster 2 rarely engaged in hands-off behavior. Furthermore, we found that individual differences were associated with the duration of hands-off occurrences but not their likelihood. These findings suggest that driver adaptation to partial automation is associated with a combination of individual factors (e.g., knowledge and experience) and the experience afforded by the driving automation's characteristics (e.g., driver alerting strategy). Together, these factors interact in ways that may affect the appropriate use of the driving automation and its overall impact on safety.**

## 1. Introduction

Partial driving assistance systems are not deigned to operate autonomously and, by definition (SAE International, 2014), require driver oversight. Drivers are responsible for monitoring the environment for objects and unexpected events, as well as the operation of driver support systems. As systems provide increased comfort and convenience, drivers have been observed disengaging from their roles and engaging in various non-driving related activities (Reagan et al., 2021). To this end, most production-level driving automation systems are accompanied by some form of driver state monitoring and support. Such systems often monitor driver engagement through steering wheel torque (e.g., Tesla Autopilot) and/or gaze or head pose detection (e.g., GM SuperCruise) to trigger alerts that are intended to enforce driver engagement.

Despite the importance of drivers' engagement, human operators often exhibit degraded monitoring performance with highly automated systems (Molloy & Parasuraman, 1996; Parasuraman & Manzey, 2010). As automation levels increase, drivers may allocate less attention to the system and traffic, affecting their readiness to resume control when needed (Victor et al., 2018).

This study examines the association between drivers' hand positions and their glance behavior using naturalistic driving data from Tesla Autopilot (AP), a hands-on partial driving automation feature recommended for highway use (Tesla Inc, 2024). AP encourages its hands-on-wheel requirement through a steering wheel torque sensing system and a range of alerts, which can escalate over time and eventually prevent the driver from using the system for the remainder of the ignition cycle if they do not respond. In addition to the relationship between drivers' hand positions and glance behavior, we examine individual differences in these behaviors. This study aims to answer three research questions (RQ):

RQ1. Is there a relationship between hands-off behavior and glance behavior, which can be a proxy for driver readiness to take-over control (TOC) from the automation?
RQ2. Do individual differences exist in the relationship between hands-off behavior and forward glances?
RQ3. How are individual differences associated with hands-off behavior?

## 2. Method

This study used data curated in a previous study (Morando et al., 2020). Only the relevant information for this study is described here.

### *2.1 Data Overview*

In this study, we analyzed 298 AP disengagements involving 19 drivers across 198 trips. For each disengagement, we annotated a 30-second driving segment surrounding the TOC (20 seconds before and 10 seconds after) for glance location and steering wheel control level annotation.

### *2.2 Glance Behavior*

Each event was annotated by coders using a frame-by-frame glance coding protocol. Glance locations were classified into nine categories. In the analysis, time percentage of glances coded as "road" within the 20-second window prior to AP disengagement was calculated as percentage of on-road glances.

### *2.3 Hands-Off Behavior*

Drivers' steering wheel control was annotated into four levels (Figure 1): high, medium, low, and none (i.e., no hands on the wheel). In the analysis, the time percentage of "none" within the 20-second window prior to AP disengagement was calculated as percentage of hands-off time. As a discrete response, if the percentage of hands-off was zero, it was coded as a "no hands-off" response. If the percentage is greater than zero, it was coded as a "hands-off" response.

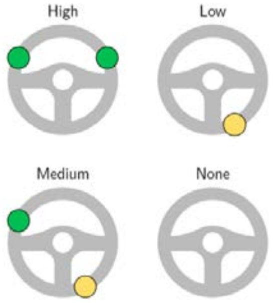



***Figure 1.*** *Hand-position coding from Morando et al., 2020*

*2.4 Analytic Methods*

To address RQ1, we fitted a generalized linear mixed-effect model to the percentage of on-road glances, with the percentage of hands-off as a predictor and the subject as a random effect.

For RQ2, we fitted a mixture model, which describes a dataset as a combination of multiple probability distributions (e.g., a mixture of different sub-populations). Based on the relationship between the percentage of on-road glances and percentage of hands-off, the model grouped drivers into two clusters. To determine differences between the two clusters, we analyzed four survey questions using the Wilcoxon test.

To address RQ3, we fitted a two-part mixed-effect model to analyze hands-off behavior as a mix of discrete response (hands-off vs. no hands-off) and continuous response (duration of hands-off). In this model, we swapped the predictor and outcome variables from the previous model, as the research question focused on hands-off behavior.

The statistical analyses were performed using R packages, with the "lme4" (Bates et al., 2015) for the mixed-effects model, the "FlexMix" (Leisch, 2004) for the mixture model, and the "brms" (Bürkner, 2017) for the two-part model.

## 3. Results

*3.1 RQ1: Is There a Relationship Between Hands-Off Behavior and Glance Behavior?*

Results showed a negative association between the percentage of hands-off state and the percentage of on-road glances (marginal $R^2$ = .07, conditional $R^2$ = .14, $p$ < .001). This suggests that drivers who spend more time with their hands off the wheel tend to spend less time looking at the road. Additionally, when examining individual data points, we observed potential bimodality in the percentage of hands-off (see Figure 2), which led our investigation in RQ2 and driver clustering.

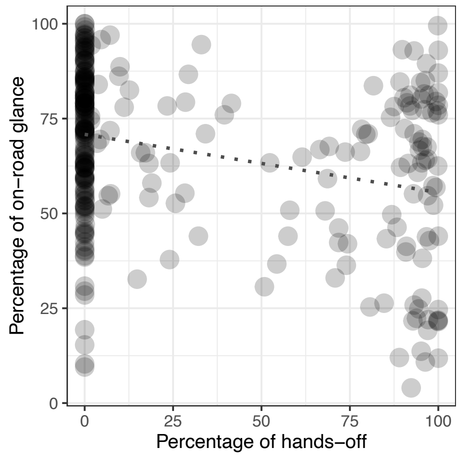



***Figure 2.*** *The relationship between the percentage of hands-off and the percentage of on-road glances showed a negative association, with the x-axis indicating potential bimodality*

*3.2 RQ2: Do Individual Differences Exist in the Relationship Between Hands-Off Behavior and Forward Glances?*

Cluster 1 comprised of nine drivers (no female, mean age = 45.8, *SD* = 13.7), while Cluster 2 consisted of 10 drivers (four female, mean age = 52.2, *SD* = 14.5). Cluster 1 drivers showed a high density of TOC events, with both no hands-off and a high percentage of hands-off, whereas Cluster 2 drivers had minimal hands-off occurrences (Figure 3). Additionally, Cluster 1 drivers tended to show a higher percentage of hands-off and lower percentage of on-road glances during AP compared to Cluster 2 drivers (Figure 4).

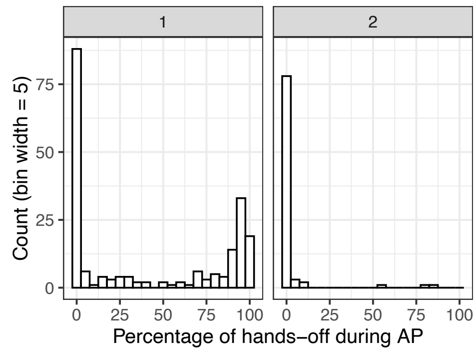



***Figure 3.*** *Cluster 1 shows bimodality, whereas Cluster 2 shows positive skewness*

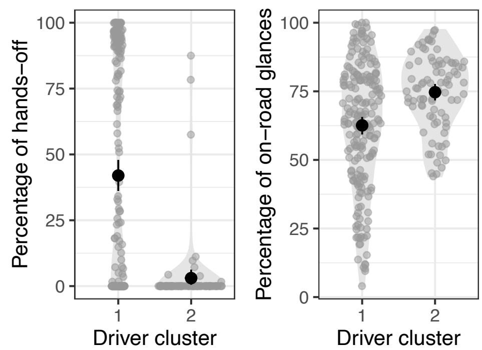



***Figure 4.*** *Cluster 1 drivers show a higher percentage of hands-off and lower percentage of on-road glances. The error bars represent 95% confidence intervals*

Survey results showed that none of the items reached $p < .05$ significance level (Table 1); however, the frequency of AP use was close to the threshold ($p = .06$). Upon qualitative examination of drivers' responses regarding AP use frequency, Cluster 1 was found to consist of more frequent AP users (Figure 5).

***Table 1*** *Wilcoxon test results showed that there were no significant differences between Cluster 1 and 2*

| Questions | *W* | *p* |
|---|---|---|
| Frequency of AP use (5-point scale) | 45 | .06 |
| Trust in technologies (10-point scale) | 26.5 | .75 |
| Trust in AP (10-point scale) | 33.5 | .46 |
| AP safety (10-point scale) | 41.5 | .16 |

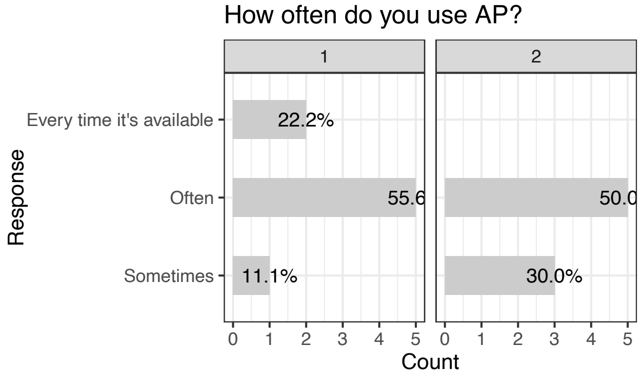



***Figure 5.*** *Cluster 1 (left) consisted of more frequent AP users compared to Cluster 2 (right)*

### *3.3 RQ3: How Are Individual Differences Associated with Hands-Off Behavior?*

Results indicated that individual differences and the percentage of on-road glances were not significantly associated with the likelihood of hands-off occurrences. However, they were significantly associated with the duration of hands-off when such occurrences happened (Figure 6). Specifically, individual differences were significantly associated with the duration of hands-off, rather than whether they occurred or not (i.e., the response).

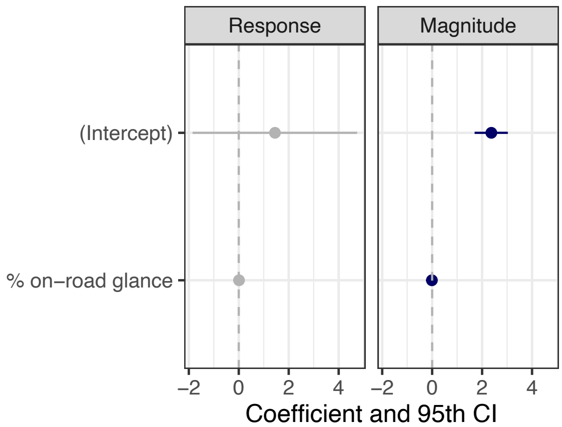



***Figure 6.*** *Model estimates and 95th percentile confidence intervals. The left column (Response) indicates hands-off events, while the right column (Magnitude) indicates the duration of hands-off periods. Dark-colored points indicate significance at the p < .05 level*

## 4. Discussion

Our findings show a negative relationship between hands-off behavior and on-road glance behavior. While further investigation is needed to understand the motivations behind these behaviors and the activities drivers engaged in during the hands-off periods, an initial review of the video data supports a link to non-driving related task engagement.

We used 20-second time windows before AP disengagement; therefore, 100% hands-off means drivers had their hands off the wheel for at least 20 seconds before the TOC. Tesla's AP features a hands-on-wheel warning system that escalates over time if the driver doesn't respond. Non-response for an extended period results in a negative outcome, such as the vehicle slowing down, and AP being locked out for the remainder of the trip. However, with this protocol, a 20-second hands-off period without a negative consequence (e.g., AP lockout) may be considered acceptable by some drivers.

In terms of safety, hands-off behavior for a partial driving automation system that requires hands-on control can be seen as a deviation from recommended safe use. However, due to the flexibility of the driver alerting strategy, drivers may not perceive their hands-off behavior as a deviation from recommended use and could normalize it over time. This may lead to suboptimal driving behavior (given limitations in system design) and an increased likelihood of adverse safety-related events. Identifying contributing factors, such as design features of monitoring and feedback systems, and developing scientifically supported mitigation strategies tailored to individual differences to enhance driver engagement while using driving automation could help drivers use these systems appropriately.

## 5. Conclusions

This study expends on previous research that separately examined hands on wheel and glance behaviors (Gershon et al., 2023; Morando et al., 2021a, 2021b). Our findings are summarized below with corresponding research questions:

- [RQ1] Drivers who spend more time with their hands off the wheel tend to spend less time looking at the road.
- [RQ2] Two driver clusters were identified: Cluster 1 engaged in both hands-off and hands-on behavior, with more frequent AP use, whereas Cluster 2 rarely engaged in hands-off behavior.
- [RQ3] Individual differences were linked to hands-off duration, not occurrence.

These findings highlight the importance of understanding and addressing individual differences in hands-off behavior to enhance the safety and effectiveness of automated driving systems.

## 6. Acknowledgments

Data for this study was drawn from work supported by the Advanced Vehicle Technologies (AVT) Consortium at MIT (http://agelab.mit.edu/avt). The analysis was supported by the Toyota Collaborative Safety Research Center.

**References**

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Gershon, P., Mehler, B., & Reimer, B. (2023). Driver response and recovery following automation initiated disengagement in real-world hands-free driving. *Traffic Injury Prevention*, *24*(4), 356–361.

Leisch, F. (2004). FlexMix: A general framework for finite mixture models and latent class regression in *R*. *Journal of Statistical Software*, *11*(8). https://doi.org/10.18637/jss.v011.i08

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors*, *38*(2), 311–322.

Morando, A., Gershon, P., Mehler, B., & Reimer, B. (2020). Driver-initiated Tesla Autopilot disengagements in naturalistic driving. *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 57–65. https://doi.org/10.1145/3409120.3410644

Morando, A., Gershon, P., Mehler, B., & Reimer, B. (2021a). A model for naturalistic glance behavior around Tesla Autopilot disengagements. *Accident Analysis and Prevention*, *161*, 106348. https://doi.org/10.1016/j.aap.2021.106348

Morando, A., Gershon, P., Mehler, B., & Reimer, B. (2021b). *Visual attention and steering wheel control: From engagement to disengagement of Tesla Autopilot*. *65*(1), 1390–1394.

Parasuraman, R., & Manzey, D. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*. https://doi.org/10.1177/0018720810376055

Reagan, I. J., Teoh, E. R., Cicchino, J. B., Gershon, P., Reimer, B., Mehler, B., & Seppelt, B. (2021). Disengagement from driving when using automation during a 4-week field trial. *Transportation Research Part F: Traffic Psychology and Behaviour*, *82*, 400–411.

SAE International. (2014). *Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems (J3016)* (pp. 1–12).

Tesla Inc. (2024). *Model S Owner's Manual* [Topic]. Tesla. https://www.tesla.com/ownersmanual/models/en_us/

Victor, T., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Aust, M. L. (2018). Automation expectation mismatch: Incorrect prediction despite eyes on threat and hands on wheel. *Human Factors*. https://doi.org/10.1177/0018720818788164

# Exploring Visual Attention Dynamics in Tram Driving. Differentiating between Expert and Novice Drivers for Gaze-Enhanced Training

Anna Warchoł-Jakubowska[1*], Krzysztof Krejtz[1], Filippo Baldisserotto[1], Iga Szwoch[1], Izabela Krejtz[1]

[1] Institute of Psychology, SWPS University, Warsaw, Poland
[*]awarchol-jakubowska@swps.edu.pl

**Abstract: The present eye-tracking study explores dynamics of visual attention and its distribution in tram drivers. While driving a tram simulator, experienced (n = 25) and novice drivers (n = 20) were focal mainly at tram stop operations. Novices switched to ambient attention faster than experts who were more attentive at the beginning of taking turns. The analysis of attention dynamics showed that experts' changes between focal and ambient attention are more relevant to the demands of the ongoing tram operation to maintain safety. Based on these findings visual attention training for novices will be designed.**

## 1. Introduction

Tram operators must be highly attentive and predictive to safely navigate urban traffic (Naweed & Rose, 2015). They oversee tram and passenger safety, maintain the tram schedule, and manage other tasks of high cognitive demand (Naweed, 2013; Naznin et al., 2017). Those tasks force drivers to maintain a highly attentive state for a long time.

Driver distraction and inattention, defined as the diversion of attention away from activities critical for safe driving towards a competing activity (Lee et al., 2009), are the main factors of vehicle road incidents (Regan et al., 2011). The Municipal Road Authority (2023) reports that most of the tram collisions in Warsaw, Poland are caused by drivers with less than 5 years of experience. There is a pressing need to train novices in maintaining attentive states for a long time and in effectively allocating visual attention to key elements of a visual field, especially at crucial tram operating tasks, tram stops operating or turns-making.

The studies on the visual attention of tram drivers compared to other road users are scarce (Kapitaniak et al., 2015). The present eye-tracking study aims to fill this gap by exploring the differences between novice and expert tram drivers in how they allocate their visual attention during challenging tasks (tram stop operations, turn-making, reactions to track hazards). The study aims to establish a foundation for attention-based training for novice tram drivers in the future.

### *1.1 Visual Attention Dynamics and Strategies of Novices and Experts*

Eye movements reflect changes in overt visual attention (Posner et al., 2004). Previous studies indicate that experts' visual strategies differ from novices across various domains (Brams et al., 2019). Warchoł-Jakubowska et al. (2023) revealed significant differences in visual attention patterns between novice and expert tram drivers while watching pre-recorded first-person tram rides. Expert tram drivers exhibited more attentive patterns of their eye movements over a longer time compared to novices. Moreover, novices tended to focus their visual attention on less relevant tram control panel elements.

The dynamics of visual attention is a constant interaction between two modes - ambient and focal. These modes refer to the difference between exploring and investigating the visual field. The distinction was first observed by Traverten (1968). The ambient mode is associated with quick and effortless processing of simple visual elements across a wide area of the visual field. The focal mode, on the other hand, requires significant cognitive resources to process complex visual information from a relatively small area of the field. Eye movements play a crucial role in this interplay of fixation duration and subsequent saccadic amplitude (Velichovsky et al., 2005).

According to Krejtz et al. (2012), ambient and focal modes are two extremes of the attention process. During attending a new scene, the ambient mode comes first and is followed by the focal mode. To measure this process dynamics, a second-order metric (Duchowski, 2017) of the attention process was introduced known as the K-coefficient (Krejtz et al., 2016). The coefficient is based on the relation between fixation duration and the subsequent saccadic amplitude, and its negative values indicate the ambient mode of attention while positive values indicate the focal mode. The absolute value of the K-coefficient determines the intensity of the current mode.

### *1.2 The Present Study*

This study attempts to determine whether expert drivers exhibit attentional patterns that can improve tram safety. Two hypotheses were developed for the study. The first hypothesis is that experts maintain more focal attention during the critical phases of different tram tasks. The second hypothesis is that experts pay more attention to instruments (control panels, windshields, or mirrors) that are more relevant to certain task phases.

To address the hypotheses, a 2x3 mixed-design experimental eye-tracking study was conducted on two groups of tram drivers (experts vs. novices) performing a test ride in the tram driving simulator. Three different tasks were embedded into the ride: hazardous track intrusion vs. tram stop operating vs. turn-taking (see Fig. 1).

## 2. Method

### *2.1 Participants*

***Fig. 1.*** *Study settings. Driver operates the simulator while wearing eye-tracking glasses*

Experts ($n$ = 25, 5 females, $M_{age}$ = 43.07, $SD$ = 5.97) and novice tram drivers ($n$ = 20, 5 females, $M_{age}$ = 37.90, $SD$ = 7.54) participated in the study. Experts had a minimum of 5 years whereas novices had a maximum of one year of accident-free tram driving experience.

### *2.2 Study Procedure and Equipment*

The study was conducted at the Warsaw Tram Training Centre using the LANDER Tram Simulator. The simulator faithfully replicated the *Pesa 120Na* tram model, and the route was based on the track infrastructure of Warsaw city in Poland.

After signing the consent form, participants were asked to perform a simulator ride on a pre-programmed route with three types of tasks: 1. potentially hazardous events (e.g., pedestrians stepping out from behind a tram, vehicle forcing right-of-way), 2. tram stops operating, and 3. right and left turns making (each turn requiring timely switch point operation). During the experimental task participants' eye movements were recorded with PupilLabs Invisible mobile eye tracker with 200Hz sampling rate.

Finally, participants rated the difficulty on a 5-point Likert-type scale and completed a demographic survey including age and gender. The first author's institutional IRB approved the study protocol (no. 54/2022).

### *2.3 Analytical plan*

To delve into the dynamics of attention, time on each task was divided into three equal time epochs (P1, P2, and P3) for each participant. Analyses of attention allocation were based on areas of interest (AOIs) windshields, control panels, and mirrors. Those three AOIs covered almost the entire visual field for each participant. The statistical analyses were based on Linear Mixed Modelling (LMM) conducted in R language for statistical computing (R Core Team, 2020).

## 3. Results

There was no statistically significant difference in the evaluation of experimental task difficulty between experts ($M$ = 2.05, $SD$ = 1.15) and novices ($M$ = 1.96, $SD$ = 0.98), $t(37.51) = 0.28$, $p$ = 0.79.

In the analysis of ambient-focal attention (see Fig. 2), the full LMM model showed that the three-way interaction significantly improved the null model, $Chi^2(4) = 9.59$, $p < .04$, *Pseudo–*$R^2$*(total)* = .08. The pairwise comparisons of estimated means revealed that while turn-taking novices were significantly more ambient when progressing with the task, from P1 to P2 (z = 3.87, $p$ < .001) and P3 ($z$ = 4.05, $p$ < .001). Experts' attention stayed more focal, there was no significant decrease between P1 and P2 ($z$ = 0.45, $p$ > .05), and P3 ($z$ = 1.50, $p$ > .05). Interestingly, during hazardous track intrusions and tram stop operation, dynamics of ambient-focal attention were similar in experts and novices (see Fig. 2).

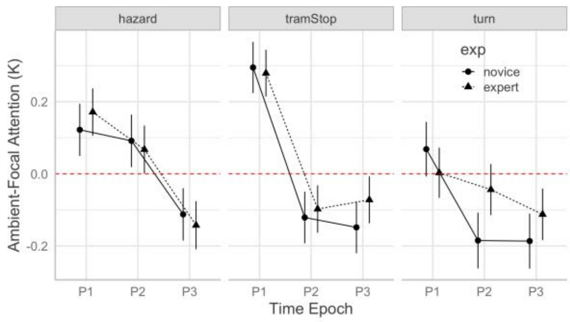



***Fig. 2****. Ambient-focal attention dynamics depending on expertise, task type, and time epochs. Note: P1, P2, P3 - time periods; error bars represent +/-1SE for estimated means*

To unveil the nature of these differences and examine the attention distribution in different periods of tasks (hazard, tram stop operating, turning), we repeated the LMM analyses separately for each category of AOIs: windshields, control panels, and mirrors with the total fixation time as a dependent variable.

The LMM analysis revealed a significant interaction effect between task, expertise, and time, $\chi^2(6) = 18.04$, $p$ < .001, *Pseudo–*$R^2$*(total)* = .51 (see Fig. 3). Pairwise comparisons indicated that experts exhibit significantly longer total fixation time on mirrors at the very first stage of turning than novices, $t(70.2) = 3.16$, $p < .01$.

Additionally, a similar analysis of the number of fixations on mirrors corroborated the findings. Experts made significantly more fixations to mirrors. Analogous analyses for windshields and control panels did not show any significant differences between experts and novices.

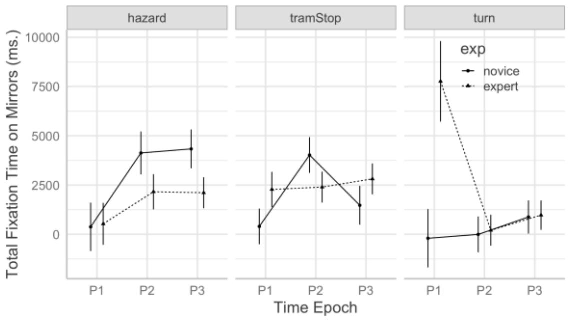



***Fig. 3.*** *Total fixation time on mirrors depending on expertise (exp), task, and time. Note: error bars represent +/-1SE for estimated means, and P1, P2, P3 represent the time periods of each task.*

## 4. Discussion and Conclusion

Literature reviews continue to indicate expertise-related differences in characteristics of visual attention, namely fixation duration, number of fixations, gaze location, and quiet eye duration (e.g., Klostermann & Moeinirad, 2019). Selective attention allocation towards relevant visual information predominates among most experts (more fixations with longer durations on important visual information), except for experts in the field of medicine (Krejtz et al. 2023). Similarly, Negi et al. (2019) demonstrated that race-car drivers (experts) adopt different driving strategies and exhibit higher steering activity resulting in lower lap-times than regular drivers in a simulated car race.

In line, the present findings reveal that tram-driving experts employ a more deliberate visual exploration strategy, likely allocating greater cognitive resources, particularly during turning. Being less attentive to tram stops and turns poses risks to passengers and nearby vehicles. Checking mirrors is essential per tram driving regulations. The focus of tram-driving experts on mirrors during turning suggests vigilance in anticipating incidents. Other studies corroborate our findings. Underwood (2007) suggests that novice drivers are relatively insensitive to changes in road traffic conditions, whereas experienced drivers anticipate potential problems by looking at parts of the road where other vehicles may intersect. But as drivers' skills develop an increase in visual scanning can be observed (Underwood, 2007).

These findings guide our focus when designing training for novices. Continuing the present work, we consider the use of augmented reality for training purposes offering interactive training. Based on our results, the training scenario would involve precise guidance on where to direct attention, such as highlighting mirrors during turns with XR glasses.

Switching from ambient to focal attention is particularly important in the context of semi-automated driving when humans need to intervene at certain moments. As drivers, especially novices, in semi-automated vehicles show poorer visual behaviour compared to active drivers in the task of transitioning from inattention to attention, (Ouddiz et al., 2020).

In our study, visual attention was measured during tram driving using a simulator. In future research, it would be valuable to investigate differences between experts and novices in real tram driving within urban environments. Additionally, it is worth noting the underrepresentation of women in tram driver roles, so future studies may explore gender disparities in tram driving.

The present dynamical approach signifies progress toward a personalized gaze-based training for novice tram drivers, akin to methods used for pilots (Muehlethaler & Knecht, 2016) and air traffic controllers (Kang & Landry, 2014). Analysing visual attention dynamics is crucial for novice training and expert insight. Training focal attention allocation aims to reduce incidents resulting from distractions during tram driving tasks.

## 5. Acknowledgments

Ministry of Education and Science Implementation Doctorate Project no. DWD/5/0498/202 financed the study. We extend our gratitude to the Warsaw Trams Management Board for granting permission to carry out the study, and to the driving instructors for their valuable assistance in preparing the research.

## References

Brams, S., Ziv, G., Levin, O., Spitz, J., Wagemans, J., Williams, A. M., & Helsen, W. F. (2019). The relationship between gaze behavior, expertise, and performance: A systematic review. *Psychological Bulletin, 145*(10), 980– 1027. http://dx.doi.org/10.1037/bul0000207

Deng, M., Wu, F., Gu, X., & Xu, L. (2021). A comparison of visual ability and its importance awareness between novice and experienced drivers. *International Journal of Industrial Ergonomics, 83*, 103141. https://doi.org/10.1016/j.ergon.2021.103141

Duchowski, T. A. (2017). Eye tracking: methodology theory and practice. Springer International Publishing AG.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological review, 87*(4), 329. https://doi.org/10.1037/0033-295X.87.4.329

Kang, Z., & Landry, S. J. (2014). Using scan paths as a learning method for a conflict detection task of multiple target tracking. *Human factors, 56*(6), 1150–1162. https://doi.org/10.1177/0018720814523066

Kapitaniak, B., Walczak, M., Kosobudzki, M., Jóźwiak, Z., & Bortkiewicz, A. (2015). Application of eye-tracking in drivers testing: A review of research. *International Journal of Occupational Medicine and Environmental Health, 28*(6), 941–954. https://doi.org/10.13075/ijomeh.1896.00317

Kieras, D. E., & Just, M. A. (Eds.). (2018). *New methods in reading comprehension research.* Routledge

Klostermann, A., & Moeinirad, S. (2019). Fewer fixations of longer duration? Expert gaze behavior revisited. German Journal of Exercise and Sport Research. 50. 1-16. 10.1007/s12662-019-00616-y

Krejtz, K., Krejtz, I., Duchowski, A., Szarkowska, A., & Walczak, A. (2012). Multimodal Learning With Audio Description: an eye-tracking study children's gaze during a visual recognition task. In *Proceedings of the ACM symposium on applied perception* (pp. 83-90).

Krejtz, K., Duchowski, A. T., Krejtz, I., Szarkowska, A., & Kopacz, A. (2016). Discerning ambient/focal attention with coefficient K. *ACM Transactions on Applied Perception, 13*(3), 11–20. https://doi.org/10.1145/2896452

Krejtz, K., Szczecinski, P., Pawlowska, A., Rutkowska-Siuda, D., Wisiecka, K., Milczarski, P., Hlobaz, A., Duchowski, A. T., & Krejtz, I. (2023). A Unified Look at Cultural Heritage: Comparison of Aggregated Scanpaths over Architectural Artifacts. *Proceedings of the ACM on Human-Computer Interaction, 7(ETRA),* 1–17. https://doi.org/10.1145/3591138

Lee, J. D., Young, K. L., & Regan, M. A. (2008). Defining driver distraction. *Driver distraction: Theory, effects, and mitigation, 13*(4), 31-40.

Muehlethaler, C.M., & Knecht, C. (2016). Situation Awareness Training for General Aviation Pilots using Eye Tracking. https://doi.org/10.1016/j.ifacol.2016.10.463

Municipal Road Authority. (2023). *Report on Road Safety in the Capital City of Warsaw.* https://zdm.waw.pl/wp-content/uploads/2023/07/Raport-ZDM-2022__online_czytniki___POP.pdf

Naweed, A. (2013). Psychological factors for driver distraction and inattention in the Australian and New Zealand rail industry. *Accident Analysis & Prevention, 60*, 193–204. https://doi.org/10.1016/j.aap.2013.08.022

Naweed, A., & Rose, J. (2015). “It's a Frightful Scenario”: A Study of Tram Collisions on a Mixed-traffic Environment in an Australian Metropolitan Setting. *Procedia Manufacturing, 3*, 2706-2713.

Naznin, F., Currie, G., & Logan, D. (2017). Key challenges in tram/streetcar driving from the tram driver’s perspective – A qualitative study. *Transportation Research Part F: Traffic Psychology and Behaviour, 49*, 39–48. https://doi.org/10.1016/j.trf.2017.06.003

Negi, N., Leeuwen, P.M. & Happee, R. (2019). Differences in Driver Behaviour between Race and Experienced Drivers: A Driving Simulator Study. 360-367. 10.5220/0007700003600367

Ouddiz, S., Paubel, P. & Lemercier, C. (2020). How do novice and expert drivers prepare for takeover when they are drivengers of a level 3 autonomous vehicle? Investigation of their visual behaviour. Le travail humain, 83, 361-378. https://doi.org/10.3917/th.834.0361

Posner, M. I., Snyder, C. R., & Solso, R. (2004). Attention and cognitive control. *Cognitive psychology: Key readings, 205*, 55-85

R Core Team (2020). A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

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

Schuhfried G. Wiener Testsystem (Vienna Test System). Mödling, Austria: Dr Gernot Schuhfried GmbH. 2011

Trevarthen C. B. (1968). Two mechanisms of vision in primates. *Psychologische Forschung, 31*, 299–337, https://doi.org/10.1007/BF00422717

Underwood, G. (2007). Visual attention and the transition from novice to advanced driver. Ergonomics, 50(8), 1235–1249. doi:10.1080/00140130701318707

Velichkovsky, B. M., Joos, M., Helmert, J. R., & Pannasch, S. (2005, July). Two Visual Systems and Their Eye Movements: Evidence from Static and Dynamic Scene Perception. In *Cogsci 2005: Proceedings of the XXVII Conference of the Cognitive Science Society* (pp. 2283–2288). Stresa, Italy

Warchoł-Jakubowska, A., Krejtz, I., & Krejtz, K. (2023). An irrelevant look of novice tram driver. In *Symposium on Eye Tracking Research and Applications* https://doi.org/10.1145/3588015.3589514

# Can the preview of an upcoming curve mitigate the effects of cognitive load on gaze patterns of expert and non-expert drivers?

Matea Ćelić [1*], Ivan Krajinović [1], Jac Billington [2], Natasha Merat [3]

[1] Rimac Technology, Sveta Nedelja, Croatia
[2] School of Psychology, University of Leeds, Leeds, UK
[3] Institute for Transport Studies, University of Leeds, Leeds, UK
[*]tsmc@leeds.ac.uk

**Abstract: It is well established that cognitive load, caused by activities which take drivers' minds off the road, leads to increased concentration of gaze toward the road centre and consequent decrease in visual anticipation of the future path, especially in the case of less experienced drivers. However, the mechanisms underlying anticipation, particularly how drivers who drive with or without a concurrent secondary activity make use of pre-event cues to anticipate the future path, are still largely unknown. This study investigated differences in gaze behaviour of cognitively loaded expert and non-expert drivers in situations with and without predictive cue which indicated curve direction and the approximate degree of the turn required. The study included 14 expert drivers (advanced police drivers) and 20 experienced, but non-expert drivers. Cognitive load- and predictive cue related differences in gaze patterns between expert and non-expert drivers will be analysed and discussed together with potential safety implications.**

## 1. Introduction

More than 50% of road traffic fatalities in the EU occur on sharp bends of rural roads, with men being three times as likely as women to die in car accidents (ERSO, 2021a). The ubiquitous addition of mobile technologies has only worsened this state in recent years, leading to increased engagement in non-driving related activities, with primarily cognitive distractions (e.g., hands-free phone conversations) accounting for almost 90% of collision incidents (ERSO, 2021b). Even though extensive research has been conducted on the impact of cognitive distraction caused by activities which increase drivers' cognitive load without redirecting their gaze from the road ahead, its safety implications, as well as potential countermeasures, are still unknown.

### *1.1 Cognitive load and driving performance*

Although there is no single accepted theory of cognitive load, the key assumption of many is the concept of a limited processing capacity, due to which identical task requirements may place different demands on novice, experienced, and highly trained experts (e.g., Engström et al., 2017; Fuller, 2005; Wickens, 1984). It is argued that the less processing resources are required for a certain activity, the less performance will be affected. Therefore, cognitive load is not only determined by performers' capabilities but also by task demands, both of which can be potentially influenced by context (Young et al., 2015). For example, Oviedo-Trespalacios and colleagues (2017) found that cognitively loaded drivers reduced their speed more while driving on sharp bends than on gently curved roads or straight segments.

To assist drivers in negotiating curves safely, many curves are preceded by a warning sign along the road, that indicates the direction and approximate degree of the turn required. Although several studies have questioned their effectiveness (Macdonald & Hoffmann, 1991; Summala & Hietamäki, 1984), there is evidence that warning signs may serve as implicit cues for unconscious or automatic vehicle control, especially needed under conditions of poor visibility (e.g., a sharp curve with a limited sight distance) (Crundall & Underwood, 2001).

Since many studies have shown that cognitively loading secondary tasks lead to increased concentration of gaze towards the road centre (Nuñes & Recarte, 2002; Reimer, 2009; Wang et al., 2014), and consequent decrease in visual anticipation, especially in the case of less experienced drivers (Lehtonen et al., 2014), it is possible that cognitively loaded drivers fail to notice on-road cues needed to safely navigate the curve. However, in one of the few studies on the topic, it has been shown that curve warnings, particularly those which emphasised the perceptual features of the curve (e.g., direction, severity of the curve) led to a reduced speed while approaching the curve, even when drivers' attention was engaged by cognitive secondary activity (Charlton, 2004). Therefore, properties of warning signs that were perceived unconsciously maintained their effectiveness despite cognitively loading concurrent activity.

### *1.2 Aim of the study*

Although several studies have shown that experienced drivers demonstrate more look-ahead fixations compared to less experienced drivers (e.g., Lehtonen et al., 2014), the mechanisms underlying anticipation, particularly how non-expert and expert drivers make use of temporal and spatial gains obtained through the recognition of pre-event cues are still largely unknown. Knowing how cognitive load affects experts' and non-experts' gaze patterns and whether curve warnings can mitigate the effects of cognitive load on curve anticipation of both expert and non-expert drivers could have valuable practical implications. This study will, therefore, investigate whether predictive cues help expert and non-expert drivers accurately anticipate the upcoming curve even when cognitively loaded.

## 2. Method

### 2.1 Participants

Two groups of male participants – expert drivers ($n_1$ = 14; $M_{age}$ = 43.14, $SD$ = 6.98) and non-expert drivers ($n_2$ = 20; $M_{age}$ = 38.05, $SD$ = 5.77) – took part in this study. Non-expert drivers drove more than 15000 kilometres per year and had a valid driving license for an average of 19.95 year ($SD$ = 6.41). The expert drivers group included UK advanced police drivers who had valid driving licenses for 24.64 years ($SD$ = 6.40), and drove regularly during their work shift. All expert drivers had completed advanced driver training and held a UK advanced driving permit for at least 5 years ($M$ = 11.79, $SD$ = 6.73).

### 2.2 Driving environment

The study was conducted at the University of Leeds Driving Simulator (UoLDS). Eye movements were recorded using a Smart Eye Pro eye tracker, consisting of two cameras placed on the dashboard in front of the driver.

Experimental drives consisted of a 10 m wide one-lane road (**Error! Reference source not found.**), with curved segments preceded and followed by a 700 m and 500 m long straight sections, respectively. All curves were 432 m long simple curves, with a radius of 137 m. Every exit tangent was followed by a 1500 m long filler section where no data was collected. Vehicle speed was limited to 70 mph (112.65 kph), and no other vehicles or objects were present during the experiment.

***Fig. 1.*** *Driving environment and a predictive cue used in the study*

### 2.3 Experimental design

The study followed a mixed model design and included one between-subject factor of Driving expertise (Experts, Non-experts) and two within-subject factors of Cue (Predictive cue, No-cue), and Cognitive load (No-task, 2-back task). Every participant needed to drive through 8 simple curves (4 left, 4 right), and 8 experimental drives, the order of which was counterbalanced among participants, included different combinations of the two within-subject factors (i.e., No-cue, No-task; No-cue, 2-back; Predictive cue, No-task; Predictive cue, 2-back).

A 1-second-long predictive cue, the purpose of which was to activate drivers' internal models of the environment, containing information about the curve direction and approximate degree of the turn required (Figure 1) was presented on the screen, 350 meters before curve entry. To ensure the visibility of the cue, it was displayed in the central road area just before the 2-back task started.

Regarding cognitive load manipulation, the baseline drive did not include any secondary task, while the auditory-verbal version of the 2-back task (Mehler et al., 2011) was used to induce high levels of cognitive load. Participants needed to retain sequences of numbers (from 0 to 9, randomly presented) in their working memory and repeat the number presented two numbers before the current one. The stimuli were presented at regular intervals of 2.25 seconds, and a voice recorder was used to record responses.

### 2.4 Procedure

This project was approved by the School of Psychology Research Ethics Committee, University of Leeds (PSYC-622). Participants were recruited using the UoLDS database and upon arrival, all of them received detailed verbal instructions about the experiment. After signing the informed consent, participants filled out a brief demographic questionnaire providing information about their age, years driving, annual mileage, and years holding an advanced driving permit. Participants then practised the 2-back task and, once ready, they moved to the simulator. The experimenter was present during the practice drive, which included segments of driving without any concurrent activity as well as segments where the 2-back task was presented. The practice drive lasted about 10 minutes, and after the experimenter left the simulator, the main drive started. The main drive consisted of eight curved segments preceded and followed by long straights, with the order of experimental conditions being counterbalanced among participants. Every experimental drive lasted approximately 20 minutes. Participants were instructed to drive as they normally would in a real-world setting and to perform the 2-back task as accurately and as quickly as possible. The complete session lasted ~50 min per participant.

## 3. Results

To better understand which parts of the curve were most affected by cognitive load, curves will be segmented into approach tangent, curve, and exit tangent, and all the metrics will be averaged within these segments. Eye-tracking data are planned to be pre-processed and analysed in the following weeks. In the case of acceptance of this paper, the authors plan to report the effects of cognitive load, curve preview, and road geometry on gaze patterns (gaze concentration and look-ahead fixations) of expert and non-expert drivers.

## 4. Discussion

The findings of this study will be discussed in terms of cognitive load-related differences in gaze patterns of expert and non-expert drivers. Also, changes in visual anticipation of future path as a result of the predictive cue, as well as its implications on road safety will be outlined.


## 5. Acknowledgements

This study was part of the collaborative PhD project between Rimac Technology and the Institute for Transport

Studies, University of Leeds. The authors would like to thank Michael Daly, Anthony Horrobin, and Albert Solernou Crusat for creating experimental scenarios and supporting data collection.

## References


Charlton, S. G. (2004). Perceptual and attentional effects on drivers' speed selection at curves. *Accident Analysis and Prevention*, *36*, 877-884.

Crundall, D. & Underwood, G. (2001). The priming function of road signs. *Transportation Research Part F: Traffic Psychology and Behaviour*, *4*, 187–200.

Engström, J., Markkula, G., Victor, T., & Merat, N. (2017). Effects of cognitive load on driving performance: The cognitive control hypothesis. *Human Factors*, *59*(5), 734-764.

European Road Safety Observatory (ERSO). (2021b). *Road safety thematic report – Speed*. Brussels: European Commission.

European Road Safety Observatory (ERSO). (2021a). *Road safety thematic report – Driver distraction*. Brussels: European Commission.

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis and Prevention*, *37*(3), 461–472.

Lehtonen, E., Lappi, O., Koirikivi, I., & Summala, H. (2014). Effect of driving experience on anticipatory look-ahead fixations in real curve driving. *Accident Analysis and Prevention, 70*, 195-208.

Macdonald, W. A., Hoffmann, E. R. (1991). Drivers' awareness of traffic sign information. *Ergonomics*, *34,* 585–612.

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.

Nuñes, L., & Recarte, M. A. (2002). Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *5*, 133–144.

Oviedo-Trespalacios, O., Haque, M., King, M., & Washington, S. (2017). Effects of road infrastructure and traffic complexity in speed adaptation behaviour of distracted drivers. *Accident Analysis and Prevention*, *101*, 67-77.

Reimer, B. (2009). Cognitive task complexity and the impact on drivers' visual tunnelling. *Transportation Research Record, 2138,* 13-19.

Summala, H., Hietamäki, J. (1984). Drivers' immediate responses to traffic signs. *Ergonomics, 27*, 205–216.

Wang, Y., Reimer, B., Dobres, J., & Mehler, B. (2014). The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand. *Transportation Research Part F: Traffic Psychology and Behaviour*, *26*, 227–237.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman and D.R. Davies (eds). *Varieties of attention*, London: Academic Press, pp. 63-102.

Young, K. L. & Salmon, P. M. (2015). Sharing the responsibility for driver distraction across road transport systems: A systems approach to the management of distracted driving. *Accident Analysis and Prevention*, *74*, 350-359.

# How younger and older drivers’ steering reversals change with cognitive distraction during both day and night-time driving

Hilkka Grahn [1*], İbrahim Öztürk [2], Natasha Merat [2]

[1] Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland
[2] Institute for Transport Studies, University of Leeds, Leeds, United Kingdom
[*]hilkka.grahn@jyu.fi

**Abstract: Distracted driving is known to contribute to undesirable driving outcomes, including near misses and crashes. While the adverse effects of driver distraction are well recognised, the modality of distraction and individual and environmental factors also matter. With this in mind, this study explored how age, lighting conditions, a Detection-Response Task (DRT) and cognitive load affect vehicle control, investigating the effects of steering wheel reversal rate (SWRR). A driving simulator study was conducted with 20 younger ($M_{Age}$ = 22.60, $SD_{Age}$ = 1.22) and 17 older ($M_{Age}$ = 65.82, $SD_{Age}$ = 3.78) drivers. Drivers completed two experimental drives (day-time and night-time), during which they were also required to complete the n-back and DRT tasks. The effect of these conditions on SWRR was examined separately for both 0.5° and 2.5° reversals. Results show an inverse change in night-time driving for both small and large reversal rates, with an increase for older drivers and a decrease for younger drivers compared to day-time driving. In addition, cognitive load was associated with fewer large reversals in the absence of DRT, whereas the presence of DRT resulted in an increase in both small and large steering reversals. The findings enhance our understanding of how driver distraction and other individual and environmental factors affect steering control.**

## 1. Introduction

Numerous studies (Caird et al., 2018; Lipovac et al., 2017) have explored the detrimental impacts of driver distraction on traffic safety. Driver distraction involves visual, manual, auditory, vocal, and cognitive elements, which can also be combined (Foley et al., 2013). Markkula and Engström (2006) state that visual and cognitive loads (or distractions) influence steering through the steering wheel reversal rate: cognitive load prompts micro corrections, while visual load (e.g., drivers briefly taking eyes off the road) results in larger corrections. However, human behaviour in driving is influenced by a range of factors, including age (Horberry et al., 2006), driving style (Rong et al., 2011), and lighting conditions (Wood, 2020). Moreover, the combined effects of these factors can further impact driving performance.

Expanding on prior research on secondary task effects on SWRR (Kountouriotis et al., 2016) and Öztürk et al.’s (2023) findings, this study explores the influence of age, lighting, DRT engagement, and cognitive load on SWRR. To account for drivers’ individual differences, we employ multilevel modelling.

## 2. Method

### *2.1 Participants, Design, and Apparatus*

The study involved 37 participants (20 younger: $M_{Age}$ = 22.60, $SD_{Age}$ = 1.22; 17 older: $M_{Age}$ = 65.82, $SD_{Age}$ = 3.78) with 10 younger and five older drivers being females. The study design was a 3 (cognitive task: no task, 1-back, 2-back) x 2 (lighting: day-time, night-time) x 2 (DRT: with DRT, without DRT) x 2 (age: younger, older) mixed factors design, with age as the only between-participant factor. The study used the University of Leeds Driving Simulator, featuring a Jaguar S-type in a 4 m spherical projection dome with a 300° projection angle and 8 degrees of freedom motion system.

### *2.2 Cognitive Task*

Participants performed an auditory n-back task (Mehler et al., 2011) with two difficulty levels (1-back, 2-back). In each ~30-second block, participants heard a random list of 10 digits at 2.25 s intervals, presented through the car’s speakers. Participants repeated the digit before (1 back) or two before (2 back) the last one heard.

### *2.3 Detection-Response Task*

Following the same procedure as Merat and Jamson (2008), a visual DRT measured the effects of a secondary task (here, n-back task) on driving performance. Following the ISO (2016) guidelines, the stimuli (a red circle) appeared randomly on the driving scene, presented every 3–5 seconds, remaining on the screen for one second. The circles were presented to the left or right of the driving scene at a vertical angle of 11° to 23° (from the forward viewpoint of drivers) and a horizontal area of 2° to 4° above the horizon. Participants were asked to look ahead, as the circles were visible in the peripheral vision. They pressed a button on the steering wheel as soon as they saw the stimuli. Each DRT block matched the n-back block duration (~30 s) and contained 7–9 stimuli.

### *2.4 Procedure*

The study was approved by the University of Leeds ethics committee (AREA 21-108). On arrival, participants received information and consent forms. First, participants practised the n-back and DRT tasks without driving, followed by practising driving in the simulator with a period of DRT and DRT plus 2-back task. For the main experiment, participants completed two drives with one of the two lighting conditions (counterbalanced). Each drive consisted of sections of 1-back, 2-back, DRT, DRT plus 1-back, and DRT plus 2-back on straight sections of a rural road. Participants

*Table 1: Multilevel model predicting small reversals*

| | | | | 95% Confidence Interval | | | |
|---|---|---|---|---|---|---|---|
| **Names** | **Effect** | **Estimate** | **SE** | **Lower** | **Upper** | **t** | **p** |
| (Intercept) | (Intercept) | 28.81 | 1.23 | 26.40 | 31.21 | 23.46 | < .001 |
| DRT | DRT present - No DRT | 0.48 | 0.31 | -0.12 | 1.07 | 1.55 | 0.121 |
| Lighting | Day - Night | -0.56 | 0.28 | -1.12 | -0.00 | -1.98 | 0.047 |
| Age group | Older - Younger | 1.91 | 2.45 | -2.90 | 6.71 | 0.78 | 0.442 |
| N-back 1 | 1-back - No n-back task | 1.57 | 0.37 | 0.85 | 2.29 | 4.28 | < .001 |
| N-back 2 | 2-back - No n-back task | 2.45 | 0.37 | 1.73 | 3.17 | 6.66 | < .001 |
| Lightning * Age group | Night – Day * Older – Younger | 3.94 | 0.57 | 2.82 | 5.05 | 6.94 | < .001 |
| DRT * N-back 1 | DRT present - No DRT * 1-back - No n-back task | 4.67 | 0.74 | 3.22 | 6.11 | 6.35 | < .001 |
| DRT * N-back 2 | DRT present - No DRT * 2-back - No n-back task | 4.08 | 0.74 | 2.64 | 5.52 | 5.55 | < .001 |

received £20 compensation after completing the study (cf. Öztürk et al., 2023).

### *2.5 Analysis*

SWRR was calculated for 0.5° and 2.5° reversals per minute, using Markkula and Engström's (2006) syntax. The 0.5° SWRR was conceptualised as small (micro) reversals, and 2.5° SWRR as large reversals (Kountouriotis et al., 2016). In the models, fixed factors include age group (younger, older), lighting (day, night), n-back task (no n-back, 1-back, 2-back), and DRT (not present, present), with each subject as a random effect. MATLAB R2020a was used for the data extraction and Jamovi 2.3.28.0 for the analysis.

## 3. Results

### *3.1 Small Reversals – 0.5°*

Model fit (Log Likelihood) was -4148.4428 and ICC of the intercept-only model was 63.6%. The model's explanatory power was notable (conditional $R^2$ = 0.69 and marginal $R^2$ = 0.05). Significant main effects (see Table 1) were observed for lighting ($p$ = .047, with night-time diminishing micro-SWRR) and n-back ($p < .001$, 1-back and 2-back increasing micro-SWRRs compared to no n-back).

Significant interactions were found between lighting and age ($p < .001$) as well as DRT and n-back tasks ($p < .001$). Compared to day-time, older drivers showed an increase ($p$ = .005), and younger drivers showed a decrease ($p < .001$) in small reversals during night-time (Figure 1).

When examining the transition from no n-back task to the 1-back task under both DRT conditions, the interaction differed (Figure 2): in the absence of DRT, there was a decrease in micro-SWRR from the no n-back condition to the 1-back condition (non-significant); while in the presence of DRT, SWRR showed a significant increase. Final model's ICC was 67.2%.

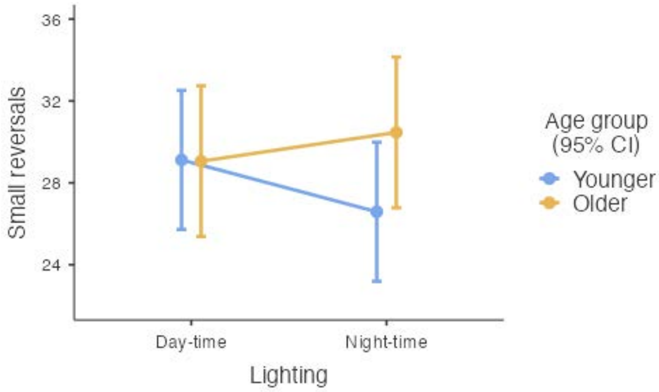



*Figure 1: Interaction of lighting and age*

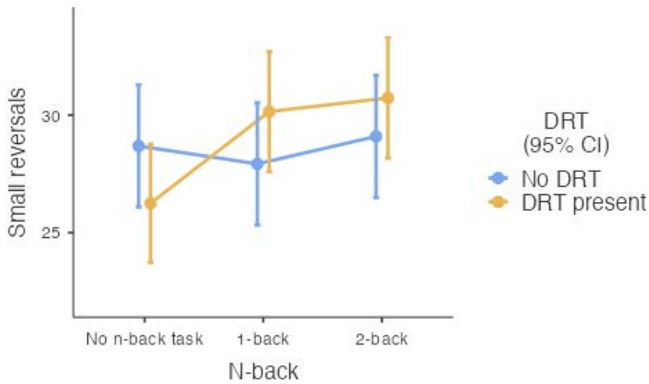



*Figure 2: Interaction of n-back tasks and DRT*

### *3.2 Large Reversals – 2.5°*

Model fit was -3168.6467 and ICC was 44.2% with conditional $R^2$ = 0.48 and marginal $R^2$ = 0.04. Significant main effects (see Table 2) were observed for n-back ($p$ = .014,

*Table 2: Multilevel model predicting large reversals*

| Names | Effect | Estimate | SE | 95% Confidence Interval Lower | Upper | t | p |
|---|---|---|---|---|---|---|---|
| (Intercept) | (Intercept) | 2.47 | 0.39 | 1.71 | 3.22 | 6.41 | < .001 |
| DRT | DRT present - No DRT | 0.36 | 0.15 | 0.07 | 0.65 | 2.46 | 0.014 |
| Lighting | Night - Day | -0.20 | 0.14 | -0.47 | 0.06 | -1.49 | 0.137 |
| Age group | Older - Younger | 0.75 | 0.77 | -0.76 | 2.26 | 0.98 | 0.335 |
| N-back 1 | 1-back - No n-back task | -0.52 | 0.18 | -0.87 | -0.17 | -2.93 | 0.003 |
| N-back 2 | 2-back - No n-back task | -0.22 | 0.18 | -0.57 | 0.13 | -1.24 | 0.217 |
| Lighting1 ✻ Age group1 | Night - Day ✻ Older - Younger | 0.87 | 0.27 | 0.34 | 1.41 | 3.19 | 0.001 |
| DRT ✻ N-back 1 | DRT present - No DRT ✻ 1-back - No n-back task | 2.88 | 0.36 | 2.18 | 3.57 | 8.11 | < .001 |
| DRT ✻ N-back 2 | DRT present - No DRT ✻ 2-back - No n-back task | 2.11 | 0.36 | 1.42 | 2.81 | 5.96 | < .001 |

1-back decreasing large SWRRs) and DRT ($p$ = .014, presence of DRT increasing large SWRRs).

Again, interactions were found between lighting and age ($p$ = .001) as well as DRT and n-back tasks ($p$ < .001). As with small reversals, night-time increased large reversals for older drivers and decreased them for younger drivers (Figure 3). The interaction (Figure 4) between n-back tasks and DRT was similar to that in the small reversals model. Final model's ICC was 45.9%.

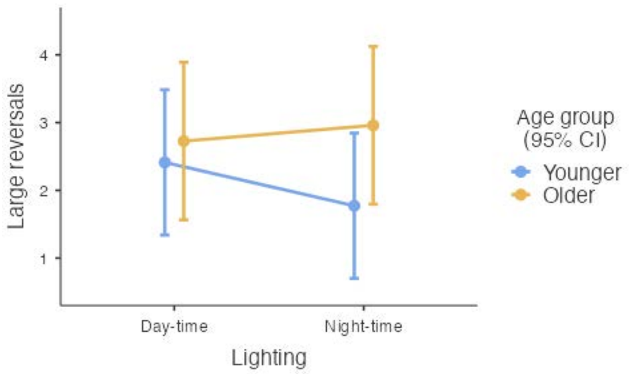



*Figure 3: Interaction of lighting and age*

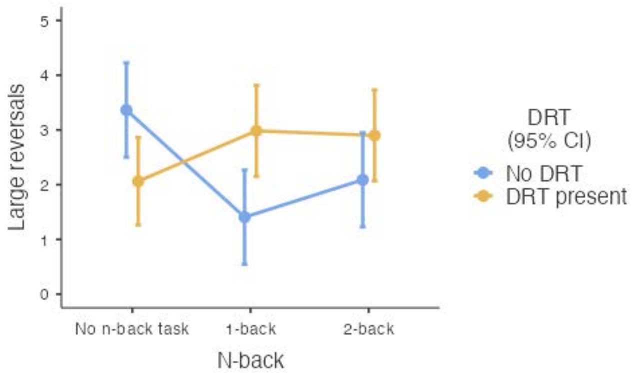



*Figure 4: Interaction of n-back tasks and DRT*

## 4. Discussion

This study investigated the effect of a number of human and environmental factors on SWRR. Without DRT, small reversals were at about the same level with increased cognitive load, but there was a significant reduction in large reversals for both the 1-back and 2-back conditions. Also, in line with Kountouriotis et al. (2016), SWRRs increased with cognitive load. For example, the increase in 2-back (compared to 1-back) might reflect the change with the increased task difficulty.

Furthermore, similar to Kountouriotis et al. (2016), larger steering wheel reversals were observed with DRT, a visual task. Despite instructions advising 'not to look (search) for the visual stimuli', the presence of the peripheral task resulted in drivers making more large reversals than when driving without DRT.

The presence of a visual task had a strong effect on large SWRRs, effectively counteracting the reduction in large reversals due to increased cognitive load. The DRT effect aligns with the Active Gaze Model (Wilkie et al., 2008), indicating that tasks diverting eyes from the road centre may result in an increase in larger steering reversals.

Previous research indicates that individual differences affect in-car glance durations (Broström et al., 2013; 2016; Grahn et al., 2023) and occlusion times (Grahn & Taipalus, 2021; Grahn et al., 2023). Here, high ICC values suggest that individual differences significantly contribute also to the variability in SWRR.

## 5. Conclusions

SWRR appears to be sensitive to individual and environmental factors as well as to different levels of cognitive load. Furthermore, the effect of visual and cognitive tasks on SWRR varies and warrants further investigation. The models also showed a large effect of individual variability in SWRR. Finally, the findings have implications for the relationship between driver distraction and driver behaviour. The change in the reversal rate of younger and older drivers during night-time driving is particularly important for road safety to understand differences in behavioural adaptation to reduced visibility during night-time driving.

**Acknowledgements**

The study was funded by the Engineering and Physical Sciences Research Council (EP/S003576/1 - HAROLD – HAzards, ROad Lighting and Driving project). Dr. Hilkka Grahn was funded by the Research Council of Finland (Appropriate Uncertainty in Manual and Automated Driving/343259) and collaborated with Dr Öztürk and Prof Merat on this work. We would like to thank Albert Solernou Crusat, Anthony Horrobin, Xian Liu, Yang Li, University of Leeds Driving Simulator Team, Prof. Richard Rowe, and Prof. Steve Fotios for their inputs throughout the study design and data collection.

 The data are available on request from the second author (I.O., i.ozturk@leeds.ac.uk).

**References**

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., 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

Foley, J. P., Young, R., Angell, L., & Domeyer, J. E. (2013, June). Towards operationalizing driver distraction. In *Driving Assessment Conference* (Vol. 7, No. 2013). University of Iowa.

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

Grahn, H., Kujala, T., Taipalus, T., Lee, J., & Lee, J. D. (2023). On the relationship between occlusion times and in-car glance durations in simulated driving. *Accident Analysis and Prevention, 182,* 106955. https://doi.org/10.1016/j.aap.2023.106955

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., & Brown, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis & Prevention, 38*(1), 185–191. https://doi.org/10.1016/j.aap.2005.09.007

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

Kountouriotis, G. K., Spyridakos, P., Carsten, O. M., & Merat, N. (2016). Identifying cognitive distraction using steering wheel reversal rates. *Accident Analysis & Prevention*, *96*, 39–45. https://doi.org/10.1016/j.aap.2016.07.032

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

Markkula,G., & Engström, J. (2006). A steering wheel reversal rate metric for assessing effects of visual and cognitive secondary task load. In Proceedings of the *13th ITS World Congress*. Leeds.

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

Merat, N., & Jamson, A. H. (2008). The Effect of Stimulus Modality on Signal Detection: Implications for Assessing the Safety of In-Vehicle Technology. *Human Factors, 50*(1), 145–158. https://doi.org/10.1518/001872008X250656

Öztürk, İ., Merat, N., Rowe, R., & Fotios, S. (2023). The effect of cognitive load on Detection-Response Task (DRT) performance during day- and night-time driving: A driving simulator study with young and older drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 97,* 155–169. https://doi.org/10.1016/j.trf.2023.07.002

Rong, J., Mao, K., & Ma, J. (2011). Effects of individual differences on driving behavior and traffic flow characteristics. *Transportation Research Record: Journal of the Transportation Research Board, 2248*(1), 1–9. https://doi.org/10.3141/2248-01

Wilkie, R. M., Wann, J. P., & Allison, R. S. (2008). Active gaze, visual look-ahead, and locomotor control. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1150–1164. https://doi.org/10.1037/0096-1523.34.5.1150

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

# Using a driving simulator study to evaluate drivers' hazard avoidance under different automation levels during day- and night-time conditions

Yee Thung Lee[1*], Rafael C. Gonçalves[1], İbrahim Öztürk[1], Karl Proctor[2], Natasha Merat[1]

[1] Institute for Transport Studies, University of Leeds, Leeds, United Kingdom
[2] Jaguar Land Rover, Coventry, United Kingdom
[*]tsytl@leeds.ac.uk

**Abstract: The aim of this study was to examine drivers' hazard avoidance under two different lighting conditions of the driving environment (daytime and night-time) and varying levels of automation (manual driving and SAE Level 2). Forty-eight participants encountered six different hazardous events during four separate drives of an urban scenario, one for each of the four conditions described above. Data analysis for this study is currently in progress, to examine how hazard avoidance (in terms of lateral and longitudinal control of the vehicle) was influenced by the different conditions. The findings of the study will be discussed in terms of their implication for the design of future driver monitoring systems, especially in automated vehicles, to understand how drivers behave when approaching hazards in a range of visibility conditions, during manual and automated driving.**

## 1. Introduction and Background

### *1.1 Hazard Perception and Night-Time Driving*

Hazard perception (HP) is an important skill to be mastered by drivers, to ensure their safety on the road. Studies have shown HP to be negatively correlated with crash involvement (Horswill et al., 2015). Due to low visibility of the driving environment, the distance at which objects become visible to drivers at night may be affected, which leads to delays in detecting hazards (Lachenmayr, 2006). Drivers also find it difficult to maintain their lane position (Öztürk et al., 2023) and have impaired longitudinal control (Yan et al., 2022) when driving at night, compared to daytime conditions. Studies have also shown compensatory behaviours from drivers at night, such as reduction of speed (Li et al., 2022; Sandberg et al., 2011) and driving closer to the road centre to avoid colliding with roadside objects (Sandberg et al., 2011).

### *1.2 Vehicle Automation*

When Partial Driving Automation is engaged (SAE Level 2; SAE, 2021), drivers can relinquish parts of the driving task to the automated system, but are required to monitor the road environment and take over manual control if needed, in order to maintain a safe drive. When steering control is managed by the Automated Driving System (ADS) drivers' perceptual-motor coupling is likely to be broken (Mole et al., 2019) as they are no longer directly involved in the driving motor control loop. Previous studies have shown that this disruption of visuomotor coordination results in higher standard deviation of lane position (Dogan et al., 2017; Madigan et al., 2017; Vogelpohl et al., 2018), lower time-to-collision (Happee et al., 2017; Radlmayr et al., 2018), and longer response to obstacle avoidance (Gold et al., 2013) after drivers resume control from automation.

### *1.3 Current Study*

While past studies on HP and night-time driving have mainly focused on performance in manual driving, how drivers negotiate around hazards in partially automated drive in such low visibility still remained unknown. In addition, most HP studies rely on simple reaction time to video-recorded events. To the best of our knowledge, there is currently a lack of understanding of drivers' hazard perception abilities in complex urban settings when Level 2 (L2) automation is engaged, and how this is affected by day and night-time lighting conditions. Therefore, the aim of this study was to investigate driver response to a range of hazards, by assessing real driving in a simulator study, comparing performance during L2 automation and manual driving conditions.

## 2. Method

### *2.1 Participants*

Forty-eight participants (20 female) aged between 25 and 63 years old (*M* = 40.31 years, *SD* = 11.04 years) were recruited. They all held a UK driving license, valid for at least three years, and were regular drivers, who drove at least once a week. Their annual mileage ranged from 3000 to 18000 miles (*M* = 8125 miles; *SD* = 3440 miles).

### *2.2 Apparatus*

The University of Leeds Driving Simulator (UoLDS), which is a high-fidelity, motion-based, driving simulator was used for this study. This consists of a Jaguar S-Type car in a 4-m spherical dome, with a 300 degrees field of view. The UoLDS' motion system is equipped with 6 degrees of freedom and allows movement for 4 cardinal directions.

### *2.3 Design*

A 2 x 2 within-subjects design was utilised with lighting of the external driving environment (Daytime and Night-Time Environment, see Fig. 1 and Fig. 2, respectively) and level of automation (L2 Automation and Manual Driving) as the independent variables. Participants completed four drives (one for each combination of conditions) in a fully counterbalanced order.

For the manual driving conditions, participants were responsible for the lateral and longitudinal control of the vehicle. They were required to adhere to the speed limit (30 mph) and drive as they would in the real world.

For the L2 automation drives, participants were required to engage the ADS whenever it was available. At the beginning of the drives, participants encountered approximately 1 minute of manual driving, followed by an auditory prompt “automation available” with the change of colour of the steering wheel icon grey to yellow. Participants were required to engage the automation as soon as possible by pressing the button on the steering wheel, or the system would automatically be engaged after 10 seconds. This turned the steering wheel icon from yellow to green. The ADS was able to maintain the vehicle’s lateral position in the centre of the lane and a driving speed of 30 mph. However, participants were told that the system was not capable of detecting and responding to on-road hazards and that they were responsible for monitoring the road environment at all times, and encouraged to take over manual control of the vehicle to maintain safety. Participants could disengage the system by turning the steering wheel, pressing the brake pedal, or pressing a button on the steering wheel. If they disengaged the system, they were required to re-engage it as soon as possible, or the system would automatically re-engage after 10 seconds. Please see Fig. 3 for the icons used for the Human Machine Interface.

***Fig. 1***. *Daytime Environment*

***Fig. 2.*** *Night-Time Environment*

### *2.4 Events and Roadway Design*

Participants drove on a two-lane urban road. The road was populated with zebra crossings, bus stops, ambient pedestrians walking on the pavement, groups of talking pedestrians, static cyclists, parked cars, and parked trucks to depict a typical urban setting. Care was taken to intersperse the position of these objects along the road, on the left and right side. A vehicle also travelled in the oncoming lane, at an average rate of one vehicle per kilometre.

Participants encountered six different events in each drive in a randomised order to account for learning effects. As per previous work on HP, each event was presented once as a materialised and once as a non-materialised incident. Materialised events were actual hazardous events that required drivers to initiate manoeuvres to avoid a collision, whereas non-materialised events were potentially hazardous events that would not develop into actual hazards even if the drivers did not reduce their speed or change their lateral position (Ventsislavova et al., 2016).

The six events were as follows: (i) a pedestrian walking along the pavement towards the road, from the left or right side of the ego vehicle (either crossing the road or stopping at the kerb – 2 materialised and 2 non-materialised events, respectively), and (ii) an oncoming car which either turned in front of the ego vehicle (materialised) or waited and turned after the ego vehicle passed (non-materialised).

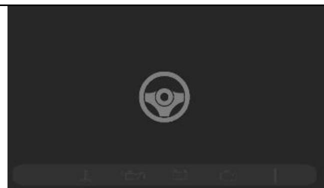
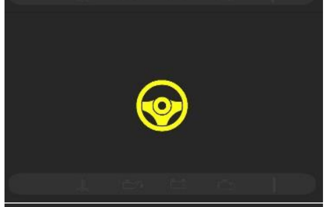
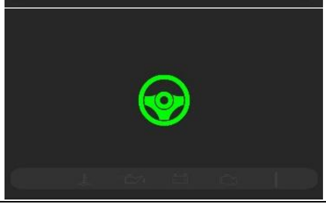
| Status | Icons |
|---|---|
| Automation Not Available |  |
| Automation Available |  |
| Automation Engaged |  |

***Fig. 3.*** *The icons used for the Human Machine Interface, placed in the dash area of the vehicle.*

### *2.5 Procedure*

The experiment was granted ethical approval by the University of Leeds Ethics Committee (BESS + FREC 2023-0792-877). After arrival, participants were briefed with the details of the experiment and provided their consent.

Participants completed two practice drives (night-time manual and daytime automation) to familiarise themselves with the driving simulator controls, and to practice how to engage and disengage the ADS. There was no hazardous event during the practice drives.

In the actual experiment, participants completed four drives. They were instructed to scan for potential on-road hazards and intervene if needed to avoid a collision. Each drive took approximately nine minutes, and participants encountered an event after approximately one minute. At the end of the study, they were compensated £30 for their time. The whole experiment took approximately 90 minutes.

## 3. Results

To date, vehicle data have been collected and are currently undergoing preprocessing and analysis. If accepted, the presentation will report on how longitudinal metrics such as mean and minimum speed, time-to-collision, and brake reaction time, as well as steering-based metrics and lateral deviation varied in response to the events, and whether these were different for manual and automated driving, during day and night-time conditions.

## 4. Discussion and Conclusions

The findings of this experiment will be discussed in terms of differences in drivers' hazard perception capabilities during day and night-time driving and how these differ between manual and L2 conditions. It is hoped that these results will provide knowledge on the design of future driver monitoring and assistance systems to help drivers with better attention management in L2 urban driving.

## 5. Acknowledgments

This driving simulator study was funded by the EPSRC HAROLD project (HAzards, ROad Lighting and Driving, EP/S003576/1). The authors would also like to thank the Simulator Team of the Institute for Transport Studies, University of Leeds, especially Albert Solernou Crusat, Anthony Horrobin and Jorge Garcia de Pedro for creating the simulated environments. The study is part of the first authors' PhD project, co-funded by Jaguar Land Rover, via an ESPRC iCASE studentship.

## References

Dogan, E., Rahal, M.-C., Deborne, R., Delhomme, P., Kemeny, A., & Perrin, J. (2017). Transition of control in a partially automated vehicle: Effects of anticipation and non-driving-related task involvement. *Transportation Research Part F: Traffic Psychology and Behaviour*, *46*, 205–215. https://doi.org/10.1016/j.trf.2017.01.012

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 1938–1942. https://doi.org/10.1177/1541931213571433

Happee, R., Gold, C., Radlmayr, J., Hergeth, S., & Bengler, K. (2017). Take-over performance in evasive manoeuvres. *Accident Analysis and Prevention*, *106*, 211–222. https://doi.org/10.1016/j.aap.2017.04.017

Horswill, M. S., Hill, A., & Wetton, M. (2015). Can a video-based hazard perception test used for driver licensing predict crash involvement? *Accident Analysis and Prevention*, *82*, 213–219. https://doi.org/10.1016/j.aap.2015.05.019

Lachenmayr, B. (2006). Traffic ophthalmology. Assessment of ability to drive in road traffic. *Der Ophthalmologe: Zeitschrift der Deutschen Ophthalmologischen Gesellschaft*, *103*(5), 425–443.

Li, G., Li, Y., Craig, B., & Liu, X. (2022). Investigating the effect of contextual factors on driving: An experimental study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *88*, 69–80. https://doi.org/10.1016/j.trf.2022.05.007

Madigan, R., Louw, T., & Merat, N. (2018). The effect of varying levels of vehicle automation on drivers' lane changing behaviour. *PLoS ONE*, *13*(2). https://doi.org/10.1371/journal.pone.0192190

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. In *Human Factors* (Vol. 61, Issue 7, pp. 1037–1065). SAGE Publications Inc. https://doi.org/10.1177/0018720819829594

Öztürk, İ., Merat, N., Rowe, R., & Fotios, S. (2023). The effect of cognitive load on Detection-Response Task (DRT) performance during day- and night-time driving: A driving simulator study with young and older drivers. *Transportation Research. Part F, Traffic Psychology and Behaviour*, *97*, 155–169. https://doi.org/10.1016/j.trf.2023.07.002

Radlmayr, J., Weinbeer, V., Löber, C., Farid, M., & Bengler, K. (2018). How automation level and system reliability influence driver performance in a cut-in situation. *Advances in Intelligent Systems and Computing*, *597*, 684–694. https://doi.org/10.1007/978-3-319-60441-1_66

SAE. (2021). Summary Of SAE International's Levels Of Driving Automation For On-ROAD Vehicles. Retrieved 29 November 2023, from https://www.sae.org/misc/pdfs/automated_driving.pdf

Sandberg, D., Anund, A., Fors, C., Kecklund, G., Karlsson, J. G., Wahde, M., & Åkerstedt, T. (2011). The characteristics of sleepiness during real driving at night - A study of driving performance, physiology and subjective experience. *Sleep*, *34*(10), 1317–1325. https://doi.org/10.5665/SLEEP.1270

Ventsislavova, P., Gugliotta, A., Penã-Suarez, E., Garcia-Fernandez, P., Eisman, E., Crundall, D., & Castro, C. (2016). What happens when drivers face hazards on the road? *Accident Analysis and Prevention*, *91*, 43–54. https://doi.org/10.1016/j.aap.2016.02.013

Vogelpohl, T., Kühn, M., Hummel, T., Gehlert, T., & Vollrath, M. (2018). Transitioning to manual driving requires additional time after automation deactivation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *55*, 464–482. https://doi.org/10.1016/j.trf.2018.03.019

Yan, Y., Zhong, S., Tian, J., & Song, L. (2022). Driving distraction at night: The impact of cell phone use on driving behaviors among young drivers. *Transportation Research. Part F, Traffic Psychology and Behaviour, 91*, 401–413. https://doi.org/10.1016/j.trf.2022.10.015

# Conflict response after assisted driving with hands on or off wheel and different steering wheel torque settings

Thomas Streubel [1*], Emma Tivesten [1,2], Mirta Zelenika Zeba [1]

[1] Volvo Cars Cooperation, Safety Center, Göteborg, Sweden
[2] Chalmers University of Technology, Göteborg, Sweden
[*] thomas.streubel@volvocars.com

**Abstract: In near-perfect assisted driving, the driver's role shifts towards supervising which may lead to reduced engagement and impaired response to unexpected events. To study this mechanism further, two test track experiments were conducted to investigate the influence of steering wheel torque settings, steering reminders, and system description on drivers' response to a conflict event after a period of assisted driving. The conflict object was a balloon vehicle placed partially in lane and appearing after a lead vehicle cut-out.**
**Reducing the system's steering wheel torque settings showed no effect on the conflict response but lowered the self-reported trust in the system. All participants with steering reminders had a normal conflict response. Some drivers without steering reminders showed no or an insufficient response to the conflict, resulting in a crash or near-crash, respectively. Most drivers with hands off the steering wheel when approaching the conflict object exhibited delayed, late or no response, despite having eyes on road.**
**While eyes on road is a good indicator of driver attention, a lack of operational control can lead to insufficient response in a conflict situation, despite knowledge of system limitations.**

## 1. Introduction

With the advancement of assisted driving, driver input is less needed. Existing SAE level 2 (Society of Automotive Engineers (SAE), 2021) systems provide longitudinal and lateral support. Recent on-market systems allow hands off wheel when driving in certain driving contexts provided that the driver is looking at the road ahead (Cantu, 2023).

Driving without being involved in operational control reduces the driver's task to supervising. A lack of driver engagement can lead to increased mind wandering (Gouraud et al., 2018), shifting attention to secondary tasks (Rudin-Brown & Parker, 2004) and impaired response to unexpected events that require the driver to act due to system limitations (Garbacik et al., 2021; Merat et al., 2012; Strand et al., 2014).

In a previous test track study, 28% of the drivers failed to avoid a crash in a conflict event after 30 minutes of supervising an otherwise highly reliable automation (Victor et al., 2018). This happened despite an explicit introduction of system's limitations and all drivers looking forward when approaching the conflict situation. Interviews revealed that some drivers expected the system to act, resulting in no or late responses to the conflict, a phenomenon known as automation expectation mismatch (Gustavsson et al., 2018; Victor et al., 2018). Interestingly, a hands on wheel requirement did not influence the outcome (Pipkorn et al., 2021), while post-crash trust in the system was higher for crashers and late responders (Gustavsson et al., 2018).

To further investigate the impact of system characteristics on driver disengagement and automation expectation mismatch, we conducted two test track studies. The drive and conflict situation remained concurrent to (Victor et al., 2018) while using a development level 2 driver assistance system with altered settings. The aim was to answer three research questions:

A. What influence have steering wheel torque settings on the drivers' trust in the system and conflict response?
B. What influence have steering reminders on drivers' conflict response?
C. What influence have the description of system capabilities and hands on/off wheel requirements on drivers' conflict response and trust during the drive?

## 2. Method

Both experiments were conducted on the test track AstaZero rural road (RISE, 2024) in Hällered, Sweden (see Figure 1). All drivers drove five laps following a lead vehicle (LV) using a development level 2 system. At the end of the drive, the participants faced a conflict situation and needed to actively intervene to avoid a crash.

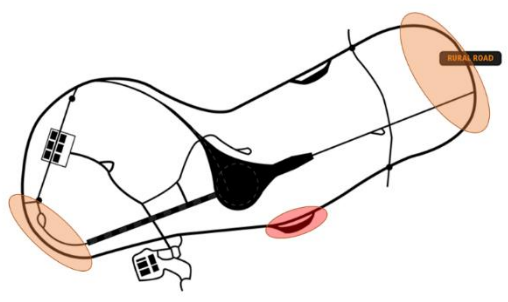



*Figure 1 AstaZero rural road including low speed areas (orange) and conflict event area (red).*

### *2.1 Participants*

In experiment 1 (E1), 56 drivers participated of which 17 were female. The age ranged between 24 and 68 years ($M$ = 41.6 years, $SD$ = 12.4 years).

Experiment 2 (E2) had 45 participants of which 13 were female. The drivers were between 26 and 62 years old ($M$ = 46 years, $SD$ = 11.5 years).

### *2.2 Test vehicle*

The test vehicle was a standard Volvo XC90 equipped with a development level 2 system similar to Pilot Assist (PA).

In E1, the steering reminders were deactivated. This inhibited a warning when the driver did not provide steering for a certain time.

In E2, the steering control was modified to create three thresholds for overriding the lane centering by the driver resulting in a low, medium, and high torque setting (see Table 1). The medium level resembles the standard PA setting, while all three levels were lower than in Victor et al (2018).

The vehicle was equipped with a logging computer and several cameras to observe the drivers' behavior.

### *2.3 Procedure*

In both studies, the drivers received a system introduction before the drive. The task was to follow a LV using the level 2 system at a set speed of 70 km/h and a time headway of 2s. The driver was fully responsible for driving and could override the system by steering or braking. The test leader was front passenger and requested a trust rating at the end of each lap.

At the end of lap 5, the conflict situation was presented as a balloon car standing partially in lane (see Figure 2). The LV executed a cut-out maneuver at about 3.3s TTC for the test vehicle (TV).

The drive was followed by a short debriefing and a post-drive questionnaire.

### *2.4 Design*

The response to the critical event is characterized by the features *eyes on road, hands on wheel* and *driver steering (DS)*. *Trust* in the system is reflected by a single item scale (0-100) based on (Lee & Moray, 1992).

In experiment 1 (E1), the participants were divided in three groups based on different introductions to the system and hands on wheel requirement. The system was introduced either as new improved Pilot Assist (PA) or as near-perfect level 2 automation (L2*) (see Table 1). In both cases, the driver was made aware of limitations and the requirement to supervise and intervene when necessary. In group 3, taking the hands off the wheel was allowed.

In experiment 2 (E2), the override threshold was varied between the three groups (see Table 1). This changed the amount of torque necessary to apply by the driver (low, medium, hard) to leave the lane center.

The conflict outcome was categorized based on lateral offset and steering onset.

*Table 1 Overview of test conditions in E1 and E2*

| **Experiment** | **E1** | | | **E2** | | |
|---|---|---|---|---|---|---|
| **Group** | **1** | **2** | **3** | **1** | **2** | **3** |
| *n* | *19* | *18* | *19* | *15* | *15* | *15* |
| Steering reminders | no | no | no | yes | yes | yes |
| Torque settings | M | M | M | L | M | H |
| System information | PA | L2* | L2* | PA | PA | PA |
| Hands on wheel requirement | yes | yes | no | yes | yes | yes |

### *2.5 Data collection and validation*

The eyes on road and hands on wheel information were acquired by annotations from videos (frame-by-frame). The onset of driver steering (DS) was determined by exceeding a steering wheel angle of 0.1 rad and a steering wheel angle speed of 0.07 rad/s. The relative position of the TV in relation to the balloon car was determined by differential GPS and vehicle sensors.

The post-questionnaire aimed to get insides in reasoning of the drivers' behavior and their expectations.

## 3. Results

This section is structured based on the consecutive steps in the conflict response followed by the conflict outcome classification and trust ratings. All times are measured in relation to reaching the conflict object.

### *3.1 Eyes on Road*

All drivers in both experiments were mainly glancing at the forward road while approaching the conflict (last 10s TTC). No differences in glance behavior between groups were found in neither experiment.

In E1, all drivers had eyes on path more than 82% of the time. The average TTC for the onset of the last eyes on road glance was 3.3s ($SD$ = 0.24s).

In E2, all drivers but one glanced at the forward road the entire time of approaching the conflict.

### *3.2 Hands on wheel*

In E1, 16 drivers had their hands off at LV cut-out start, and most of them (12) were in group 3. Four drivers were

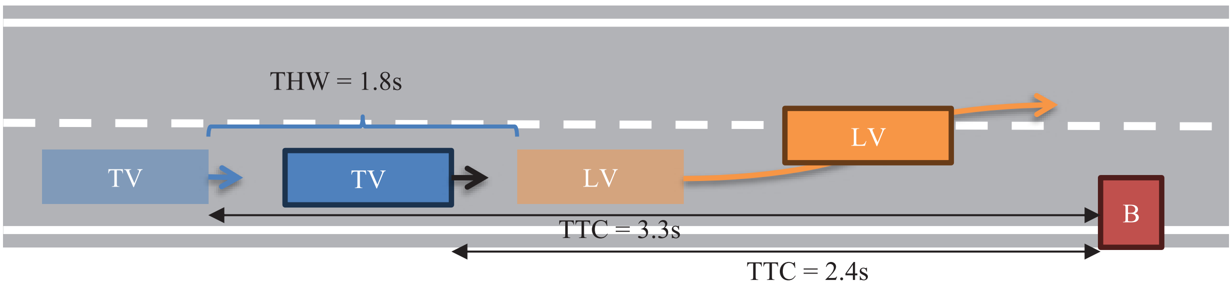



*Figure 2 Illustration of conflict event (bird-view) at start of lead vehicle (LV) cut-out (shaded) and at full reveal of balloon car (B) at 2.4s TTC for the test vehicle (TV) (opaque)*

hovering over the steering wheel, one in group 1 and three in group 2 despite hands on wheel requirement.

The mean time for hands on wheel onset was 1.9s TTC (*SD* = 0.83s) for 14 drivers that had their hands off. Two drivers kept their hands off resulting in a crash.

In E2, all drivers had their hands on the wheel during the conflict approach.

### *3.3 Driver steering*

All drivers in E2 initiated steering before the conflict, while in E1 five drivers did not steer (two each in group 1 and 3, and one in group 2). The driver steering onset per group is displayed in Figure 3.

A one-way ANOVA revealed that the drivers in the groups of E1 and E2 who steered did not differ in steering onset ($F(5,88) = 1.18$, $p = 0.33$).

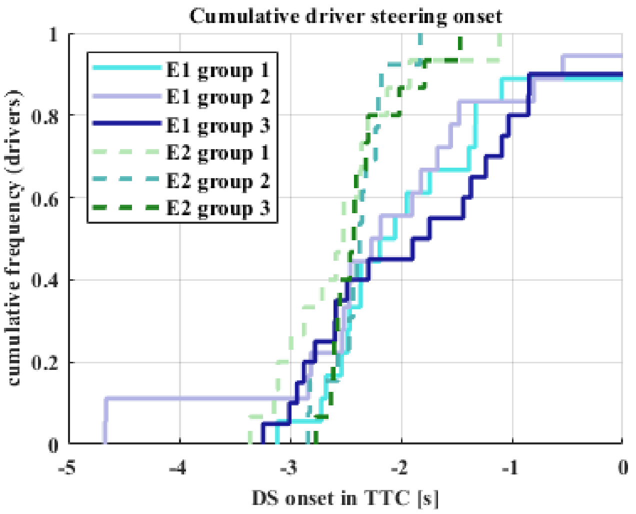



*Figure 3 Cumulative distribution of steering onset (DS)*

### *3.4 Conflict Outcome (Clustering)*

Different driver responses led to various conflict outcomes in E1. These are clustered based on lateral offset when passing the conflict object and time of driver steering in TTC (see Figure 4). In E1, six conflict events resulted in a crash due to no or too late response by the driver. Another six drivers experienced a near-crash through an insufficient response by either steering late (TTC < 1.0s) or by creating a very low lateral offset (< 0.5m). Slightly higher lateral offsets with delayed responses (TTC < 1.5s) were characterized as incidents (8 cases). The remaining drivers showed normal responses resulting in a comfortable evasion of the conflict object.

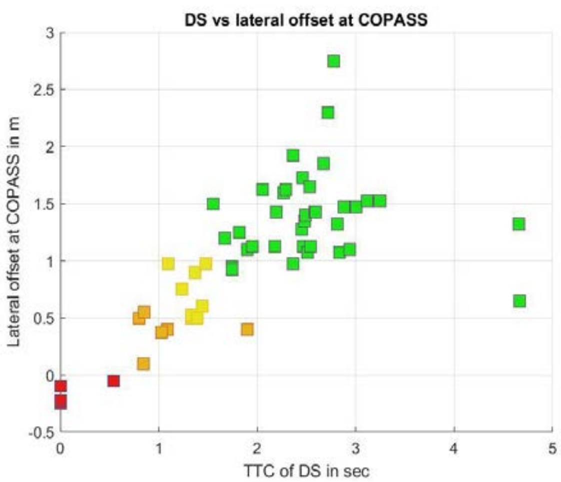



*Figure 4 E1 clustering conflict outcome based on response resulting in no or too late (red), insufficient (orange), delayed (yellow) and normal response (green)*

In E2, only one driver showed a delayed response, while the rest had a normal response.

### *3.5 Trust rating*

In E1, the trust rating was high across the three groups. There was no statistical difference between the groups. The post drive trust rating was generally lower than during the drive.

In E2, the self-reported trust is statistically significant lower in group 1 compared to both other groups for all rating points except after lap 1 (cp Figure 5).

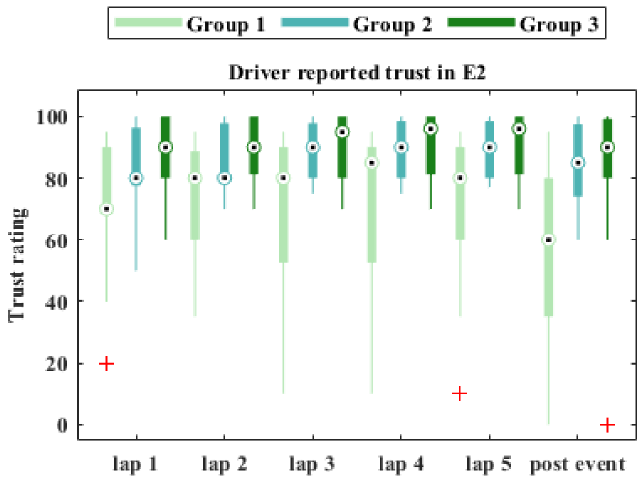



*Figure 5 Trust rating in experiment 2 across all laps with outliers (red markers)*

### *3.6 Driver expectation*

The post-drive questionnaire in E1 showed that all participants who did not respond or responded late expected the system to handle the conflict. That was also true for approximately 60% of participants who had a near-crash or a delayed response. On the other hand, only 13% of the normal responders expected a reaction from the car.

## 4. Conclusion

Different torque levels to override the lane centering functionality did not show any effect on the response to the conflict. All driver's avoided a crash by steering in a timely manner. However, the lower torque setting had a negative influence on the driver's trust in the system.

Steering reminders prevented hands off driving during the conflict approach. Of the drivers who did not receive steering reminders, several had an insufficient, late or no response to the conflict.

The hands on wheel requirement did not necessarily prevent hands off driving. In return, allowing hands off driving led to delayed responses to the conflict, requiring additional time before steering onset. The system description alone had no influence on neither trust nor conflict response nor conflict outcome. High system performance during the drive likely had a stronger influence on the driver's expectation of the system to resolve conflicts.

## 5. Acknowledgments

The authors would like to thank their colleagues Ruben Brau and Patrik Andersson for supporting the study. Furthermore, we appreciate the annotation efforts made by Lovisa Högberg, Jennifer Nordqvist and Rickard Nilsson.

**References**

Cantu, M. (2023). Which car makers offer hands-free driving systems for 2023? *The Seatle Times*. https://www.seattletimes.com/business/edmunds-who-offers-hands-free-driving-systems-for-2023/

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

Gouraud, J., Delorme, A., & Berberian, B. (2018). Influence of automation on mind wandering frequency in sustained attention. *Consciousness and Cognition*, *66*, 54–64. https://doi.org/10.1016/J.CONCOG.2018.09.012

Gustavsson, P., Victor, T. W., Johansson, J., Tivesten, E., Johansson, R., & Aust, L. (2018). What were they thinking ? Subjective experiences associated with automation expectation mismatch. *Proceedings of the 6th Driver Distraction and Inattention Conference*, *October*, 1–12. https://www.researchgate.net/publication/328388914_What_were_they_thinking_Subjective_experiences_associated_with_automation_expectation_mismatch

Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, *35*(10), 1243–1270. https://doi.org/10.1080/00140139208967392

Merat, N., Jamson, A. H., Lai, F. C. H., & Carsten, O. (2012). Highly automated driving, secondary task performance, and driver state. *Human Factors*, *54*(5), 762–771. https://doi.org/10.1177/0018720812442087

Pipkorn, L., Victor, T. W., Dozza, M., & Tivesten, E. (2021). Driver conflict response during supervised automation: Do hands on wheel matter? *Transportation Research Part F: Traffic Psychology and Behaviour*, *76*, 14–25. https://doi.org/10.1016/j.trf.2020.10.001

RISE. (2024). *AstaZero*. https://www.astazero.com/

Rudin-Brown, C. M., & Parker, H. A. (2004). Behavioural adaptation to adaptive cruise control (ACC): Implications for preventive strategies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *7*(2), 59–76. https://doi.org/10.1016/j.trf.2004.02.001

Society of Automotive Engineers (SAE). (2021). *Taxanomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*.

Strand, N., Nilsson, J., Karlsson, I. C. M. A., & Nilsson, L. (2014). Semi-automated versus highly automated driving in critical situations caused by automation failures. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*(PB), 218–228. https://doi.org/10.1016/j.trf.2014.04.005

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

# Driver Adaptation to Lower Levels of Automation (L2) Using Naturalistic Driving Data: How Long Before Drivers Start Looking Away from the Forward Roadway?

Sheila G Klauer[1a], Gabrial Anderson[1*], Shu Han[1], Feng Guo[1], and Thomas Fincannon[2]

[1] Virginia Tech Transportation Institute, 3500 Transportation Research Plaza, Blacksburg, VA 24060, USA
[2] National Highway Traffic Safety Administration, 1200 New Jersey Avenue, SE, Washington, DC 20590, USA
[*]cklauer@vtti.vt.edu

**Abstract: This project evaluated driver adaptation in the hours, days, and months after the introduction of Level 2 (L2) advanced driver assistance system features into the driving task. Existing naturalistic driving study databases were used for analysis. To assess driver adaptation, the analysis identified three phases of exposure time to L2 features: Phase 1 (under 3 hours), Phase 2 (3 hours up to 8 hours), and Phase 3 (over 8 hours). The results suggested that driver adaptation was present for high-risk secondary tasks, as significant increases in engagement were observed over the three phases, but only when L2 features were active. Additionally, drivers set their vehicle speed above the speed limit more frequently between Phases 1 and 2, with higher speeds set when L2 features were active as opposed to when they were inactive. While these results may be concerning, larger scale efforts are needed to determine if increased crash risk associated with speeding and high-risk secondary task engagement exists with L2 features active. Additionally, we need to better understand the impact of traffic/roadway conditions on speed selection with L2 systems.**

## 1. Introduction

In the driving domain, the concept of behavioral adaptation refers to how humans respond, either intentionally or unintentionally, to the introduction of a new technology that serves a specific driver need (Manser et al., 2005). More general theories of behavioral adaptation focus on risk-based measures and integration of the theories of risk compensation and risk homeostasis (Kulmala & Rämä, 2013; Taylor, 1964).

Risk allostasis theory builds on the risk homeostasis theory to incorporate driver perception and decision-making along with the constant changes that occur in the environment (i.e., learning over time). Kinnear and Helman (2010) utilized risk allostasis theory to evaluate and potentially predict behavioral adaptation for drivers using driver assistance technologies (SAE J3016). They maintained that with sufficient sensitivity to risk, task difficulty, and workload, risk allostasis theory predicts that any alteration of the driving task (i.e., introduction of advanced driver assistance systems [ADAS]) will result in driver adaptation that will trend toward maintaining task demand within a preferred range. In other words, as the automated driving features simplify the driving task, the driver will feel free to increase task demand in a variety of ways that could include increased speed, increased secondary task engagement, and decreased following distance. Thus, risk allostasis theory predicts that the use of driver assistance technologies would result in increased trust and reliance on them. This claim was substantiated by Llaneras et al. (2013), who showed that, when given the opportunity to relinquish control of lateral and longitudinal operations to a simple but reliable automated system, most drivers will engage in moderate to complex secondary tasks and will also exhibit increased eyes-off-the-forward-road time.

While behavioral adaptation can occur as a result of changes to any aspect of the roadway system, the present study is concerned specifically with how drivers initially adapt their behaviors to Level 2 (L2) ADAS features, as defined by SAE International (SAE J3016, 2018). SAE J3016 defines L2 features as the combination of both lateral and longitudinal control system support for the driver (i.e., steering and acceleration/braking). Using frequency distributions generated from the naturalistic driving databases used for these analyses, this study operationally defined these three phases of exposure to L2 system features as:

- Phase 1: under 3 hours of L2 experience;
- Phase 2: between 3 and 8 hours of L2 experience; and
- Phase 3: over 8 hours of L2 system experience.

The following research questions were answered by this analysis:

1) Do drivers look away from the roadway longer while using their L2 features?
2) How does eyes-off-road time vary by exposure phase?

## 2. Method

### *2.1 Naturalistic Databases*

This analysis utilized naturalistic driving study (NDS) databases of middle-aged drivers and compared behavior during the first weeks of driving with L2 systems (Novice L2 NDS; participants ages 25 to 54).

### *2.2 Participants*

In the Novice L2 NDS, 120 participants were recruited from the Washington, DC, metro area and drove one of 10 instrumented vehicles equipped with market-available lateral and longitudinal driving feature systems. These instrumented vehicles were "on loan" to the participants for the duration of the data collection period. Sixty-six of these participants met the qualifications for inclusion in the data analysis.

### *2.3 Independent Variables*

For the independent variables, the variable L2 system status was split into two critical states: L2 system active and

L2 system available but inactive. These two periods of driving time represented: (1) periods when L2 features were actively engaged, and (2) periods when L2 features were available to the driver but the driver chose to not engage them. These were the only two periods of driving time evaluated in these analyses.

Additionally, L2 exposure phase was assessed, where the duration of exposure to L2 systems was evaluated for all novice L2 drivers using frequency distributions. Based upon these distributions, exposure to L2 systems was split into three groups: less than 3 hours of driving with L2 systems active, greater than 3 hours but less than 8 hours of driving with L2 systems active, and 8 hours or more of driving with L2 systems active.

### *2.4 Dependent Variables*

A random sample of cases when L2 features were active was identified and reviewed by trained coders. Once the cases were identified, a matched sample of controls was identified when the L2 features were available but inactive. Each case and each matched control were 15 seconds in duration. The sampling strategy for these matched samples is described below. The key data obtained from these matched samples included specific driver behaviors such as types of secondary task engagement and 15-second epochs of continuous eye-glance location coding.

### *2.5 Sampling of Cases and Controls*

The VTTI team implemented a sampling plan designed to determine if there is driver adaptation over time while using L2 features. Given that we are interested in identifying learning over time, the sampling plan was designed to best assess the earliest moments of exposure to the L2 systems (i.e., steepest part of a prototypical learning curve). This sampling plan was designed to over-sample the steepest section of this learning curve to evaluate how driver behavior changes during this critical time.

The sampling strategy adopted for the Novice L2 NDS dataset resulted in a final matched sample (when L2 systems were active) that frequently sampled (four samples per hour) for Phase 1, moderately sampled (two samples per hour) for Phase 2, and less frequently sampled (one sample per hour) for Phase 3. The control segments (when L2 systems were available but inactive) were identified and followed the same sampling strategy (matching criteria described below). Cumulative exposure to L2 systems was used because each participant had a different number of hours of exposure when L2 systems were active.

## 3. Results

Glance metrics were calculated using the matched sample dataset. Driver glance behavior was assessed by L2 system status and L2 exposure phase. Exploratory analyses were conducted to assess the frequency of any glance away from the forward roadway by matched sample, which includes both the case sample with L2 active as well as the matched control sample with L2 available but inactive. Figure 1 shows the histogram of matched samples by total eyes-off-road duration (seconds). Over half the matched samples had no eyes-off-road time. A mixed-effects Poisson regression was used to assess significant differences in eyes-off-road count by L2 status and exposure phase, with participants used as a random effect variable (see Table 1). The results indicated a significant main effect for L2 status in that participants were looking away in more matched samples when L2 systems were active than when L2 systems were available but inactive. Additionally, there was a main effect of phase in that drivers looked away more often in Phase 3 than in Phase 1, but there was no significant difference between Phase 2 and Phase 1.

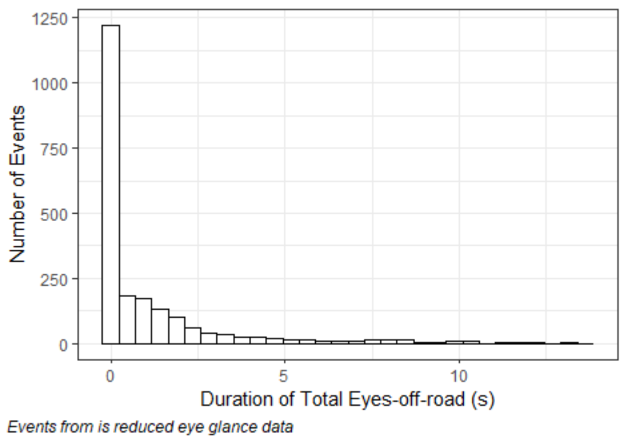



***Fig 1.*** *Matched samples distribution of total eyes-off-road time.*

***Table 1***. *Results of the mixed effects Poisson regression comparing eyes-off-road glance behavior by L2 status and exposure phase*

| | Estimated coefficient | Standard Error | *Z* value | *p* |
|---|---|---|---|---|
| Intercept | -0.065 | 0.081 | -0.81 | 0.42 |
| L2 Status (Available but inactive vs. Active) | -0.09 | -0.04 | -2.24 | 0.025 |
| Phase 2 vs. Phase 1 | 0.08 | 0.05 | 1.83 | 0.07 |
| Phase 3 vs. Phase 1 | 0.167 | 0.07 | 2.55 | 0.011 |
| Phase 3 vs. Phase 2 | 0.079 | 0.068 | 1.16 | 0.246 |

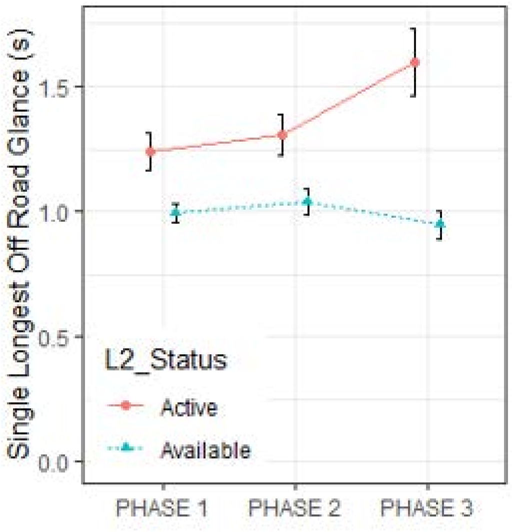



***Fig. 2**. Interaction between L2 status and exposure phase on single longest off-road glance.*

Analysis of variance (ANOVA) was conducted to assess whether eyes-off-road glance metrics were significantly longer when L2 systems were active versus available but inactive and if eye-glance durations changed over time. Four metrics—the total eyes-off-road time, the mean duration of glances, the single longest glance, and the number of glances—were computed for eyes-off-road glances. For brevity of this abstract, only single longest glance is presented in Figure 2. The results indicate that drivers looked away from the forward roadway significantly longer when the L2 system was active versus when the L2 system was available but inactive for total glance duration, $F(1, 934) = 19.30, p < 0.001$; mean glance duration, $F(1, 934) = 29.67, p < 0.001$; and single longest glance, $F(1, 934) = 31.28, p < 0.001$. The numbers of glances were not statistically different from each other, $F(1,934) = 3.03, p = 0.082$. The interaction of L2 status by phase was significant, where glance duration off road increased across phase when L2 systems were active but not when available but inactive. This result was found for total glance duration, $F(2, 934) = 8.65, p < 0.001$; mean glance duration, $F(2, 934) = 3.51, p = 0.03$; single longest glance, $F(2, 934) = 2.36, p = 0.04$; and number of glances, $F(2, 934) = 5.14, p = 0.006$.

## 4. Conclusions

As for driver behavior with L2 systems over time, we observed statistically significant increases in mean off-road glance duration, single longest off-road glance duration, and percentage of eyes-off-road time for drivers using L2 systems. We also observed a statistically significant increase in eyes-off-road time when L2 systems were active across exposure phases, in which off-road glances were shorter in Phase 1 than they were in Phase 3. These longer glances may demonstrate the presence of driver adaptation to L2 systems in that within 8 hours of exposure to L2 systems, drivers are engaging in significantly longer eyes-off-road time than they were after initial introduction to the L2 system. While these durations of having eyes off road do not approach the 2-second duration where there would also be safety concerns (Klauer et al., 2006), these longer durations may indicate that drivers are more comfortable and will engage in longer durations of eyes-off-road time, thus diminishing their ability to perform the dynamic driving task required when using these systems. Designers of L2 systems should consider the use of driver monitoring devices that incorporate driver eyes as well as head position in the algorithms for when drivers should use L2 systems to maintain safety and minimize the opportunities for unintended consequences of the deployment of this technology.

**References**

Klauer, S. G., Guo, F., Sudweeks, J. D., & Dingus, T. A. (2010). *An analysis of driver inattention using a case-crossover approach on 100-Car data* (Report No. DOT HS 811 334). National Highway Traffic Safety Administration.

Kulmala, R., & Rämä, P. (2013). Definition of behavioural adaptation. In C. Rudin-Brown & S. Jamson (Eds.), *Behavioural adaptation and road safety* (pp. 18-36). CRC Press.

Kinnear, N. A. D. & Helman, S. (2013). Updating risk allostasis theory to better understand behavioral adaptation. In C. Rudin-Brown & S. Jamson (Eds.), *Behavioural adaptation and road safety.* Taylor and Francis Group.

Manser, M. P., Ward, N. J., Kuge, N., & Boer, E. (2005). Driver behavioral adaptation in response to a novel haptic driver support system. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, San Diego, CA.

Llaneras, R. E., Salinger, J., & Green, C. A. (2013). Human factors issues associated with limited ability autonomous driving systems: Drivers' allocation of visual attention to the forward roadway. *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*. https://pdfs.semanticscholar.org/344d/7759e45e059b9b73a0de25e7d8b4af52f5bb.pdf

SAE International. (2018). *J3016 - Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*.

Summala, H. (1988). Risk control is not risk adjustment: The zero-risk theory of driver behaviour and its implications. *Ergonomics*. https://doi.org/10.1080/00140138808966694

Taylor, D. H. (1964). Drivers' galvanic skin response and the risk of accident. *Ergonomics, 7*(4), 439-451.

# Assessing Situation Awareness while Driving with Automation

Barbara Metz [1*], Johanna Wörle[1 2], Myriam Metzulat[1]

[1] WIVW GmbH, Robert-Bosch-Str. 4, 97209 Veitshöchheim, Germany
[2] Current affiliation: Singapore-ETH Centre, 1 Create Way CREATE Tower, Singapore 138602

[*]metz@wivw.de

**Abstract: In a driving simulator study, different levels of situation awareness (SA) were created experimentally and the suitability of different methods for assessing SA and their relation to each other were investigated. N=41 participants experienced driving with an automated driving function (AD) in four conditions (SAE level 2 vs. two versions of SAE level 3 vs. SAE level 3 with black scenery during AD mode) which were expected to result in different levels of SA. Results from the post- drive questionnaire indicate that while being in AD mode, three levels of SA could be induced for aspects of SA perceived via visual perception of the road. For other aspects of SA and perception during takeover request, results show reduced variation of SA between conditions with only two levels remaining. These two levels are also in-line with reaction times in takeover situations. In summary, we could induce different levels of SA and the used post-drive questionnaire was suited to measure those differences.**

## 1. Introduction

Situation awareness (SA) was first investigated in aviation and is defined as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" (Endsley, 1988, p. 97). The driver is required to maintain SA all the time while driving with level 2 AD (L2-AD; SAE, 2021) although SA might be compromised by vigilance decrements. Conversely, with level 3 AD (L3-AD; SAE, 2021) the driver is allowed to lose SA in automated mode and then needs to regain it only at takeover requests (TOR).

To measure SA, interview or questionnaire methods like SART (Taylor, 2017) are used as well as eye tracking (Moore & Gugerty, 2010; van de Merwe, van Dijk, & Zon, 2012), physiological measures (Zhang et al., 2020) or probe measures. In probe measures which are similar to SAGAT (Endsley, 1995), knowledge on the situation is asked from the subject directly in or after the situation (Strybel et al., 2016). In AD, SA is mainly investigated by means of eye tracking measures, i.e. visual attention (Liang et al., 2021), by the response to critical events (Gold et al., 2013; Merat & Jamson, 2009) and also by probe measures (e.g., Sirkin, Martelaro, Johns, & Ju, 2017).

For driving with AD, it proved difficult to show a relation between different measures of SA like SAGAT, questionnaires or takeover performance (e.g. van den Beukel, van der Voort & Eger, 2016; Cortens, 2019; Schwindt et al. 2023). However, for future research on SA in AD, it needs to be understood, which of the methods used in the literature are capable of measuring SA in AD and how the different methods relate to each other.

## 2. Method

### 2.1 *Experimental approach*

A driving simulator study was conducted with the aim to create different level of SA while driving with AD. Four conditions were implemented.

- L2: drivers are instructed to stay attentive and monitor the automation all the time.
- L3: drivers are allowed to be inattentive during the automated drive. They watch a video while being in automated mode.
- L3+: same as the L3 condition but with an extended HMI-version which provides more information in case of a TOR.
- Black: same as the L3 condition but in automated mode the visual input to the driver is reduced by not showing the driving environment. During AD mode, the projected scenery is turned black. The scenery becomes visible at the beginning of a TOR and remains visible during manual driving.

Every driver participated in two experimental sessions of about 3 hours each and experienced all four conditions in randomized order.

The study took place in the high-fidelity moving base driving simulator of the WIVW GmbH. The simulation software was SILAB® Version 6.0 (WIVW GmbH, Veitshöchheim). In the study, participants drove with a simulated L2/L3 motorway AD which could automatically adjust speed and distance, stay in its lane and overtake slower vehicles. In case of a system boundary, control was given back after an acoustical and visual TOR with a total takeover time of 10 seconds. At a TOR drivers had to deactivate the AD by button press and continue driving manually.

The AD was experienced during drives of 50 minutes, in which sections with stable driving in AD alternated with situations with system boundaries. Per drive, between 7 and 8 TORs occurred in situations with mostly medium demands to the driver like construction sites or highway intersections.

### 2.1 *Data collection*

Data collection consisted of a variety of different measures for assessing SA:

- a post-drive questionnaire assessing the perception of different aspects of the driving situation while being in automated mode and at TORs using a 5-point Likert-scale

- experienced situational criticality directly after each TOR on a 10-point scale (based on Neukum et al. (2008), range: 1=harmless, 10=uncontrollable)
- logging of vehicle and AD state to calculate metrices like takeover reaction times to TORs.

### 2.2 *Sample*

The study was conducted with N=41 (20 female, 21 male) participants. On average they were 45 years of age (sd=16.3).

### 2.3 *Analyses*

In the following, the focus is on the analysis of subjective data. Takeover performance is assessed by analysing the reaction time until the AD system is deactivated. For measures collected per takeover situation, indicators are first average per driver and condition and are then analysed.

## 3. Results

### 3.1 *Post-drive evaluation*

In the post-drive evaluation (see figure 1a), there were significant effects of condition (see table 1) on items asking for situational knowledge while driving in automated mode. For items focusing on the visual perception of relevant aspects of the driving scenery (lane, other vehicles etc.) the highest ratings were given in the L2, the lowest in the black condition with the two L3 conditions laying in between. For items relating to aspects that can be perceived outside the forward view (speed on speedometer, vehicle dynamics, sounds) the difference lies between L2 and the other three conditions.

***Table 1*** *Results of ANOVAs on impact of condition on subjective evaluation.*

| Item | df | F | p |
|---|---|---|---|
| knew in which lane | 3,111 | 32.6 | <.001 |
| knew car in front | 3,111 | 52.3 | <.001 |
| knew my speed | 3,111 | 18.9 | <.001 |
| focused on driving | 3,111 | 47.8 | <.001 |
| observed environment | 3,111 | 64.6 | <.001 |
| were aware of sounds | 3,111 | 7.9 | <.001 |
| knew car on next lane | 3,111 | 68.7 | <.001 |
| focused on other things | 3,111 | 35.5 | <.001 |
| felt vehicle dynamics | 3,111 | 4.2 | <.01 |
| knew why TOR | 3,111 | 8.5 | <.001 |
| knew on which lane | 3,111 | 7.4 | <.001 |
| knew vehicles around me | 3,111 | 8.4 | <.001 |
| knew how to take control back | 3,111 | 3.4 | <.05 |

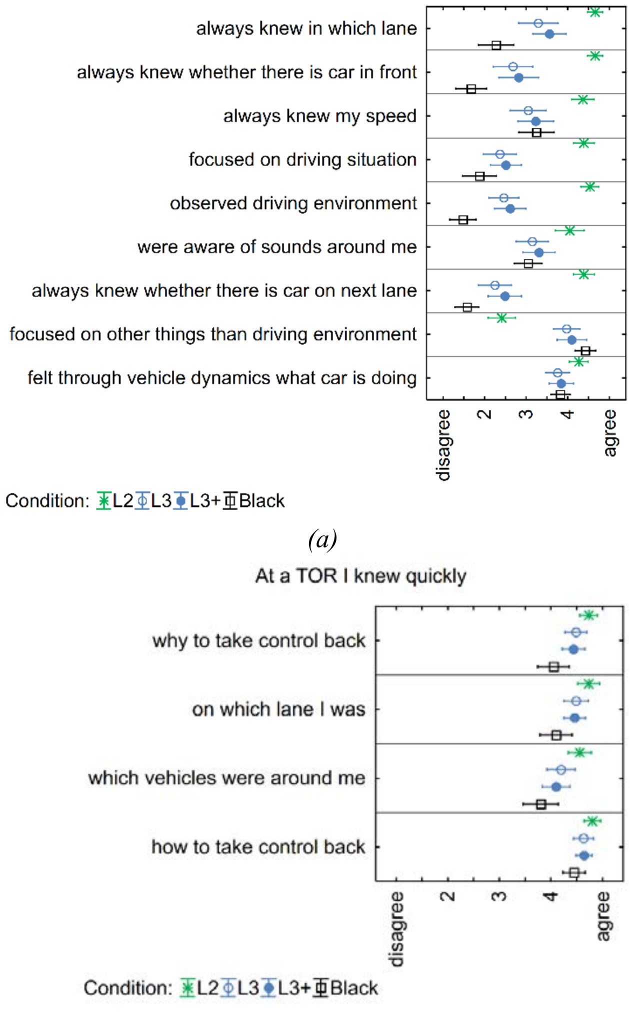



***Figure 1.*** *Post-drive ratings on situational knowledge and attention (a) while being in AD mode (b) during TORs.*

If it comes to assessing the driving situation at the time of a TOR (figure 1b), participants highly agreed in all conditions that they knew quickly important aspects of the situation and what to do. Again, there was a significant impact of condition (see table 1). Subjects reported a significantly quicker perception of the situation in the L2 compared to the black condition for all items, with L3 and L3+ lying in between.

Average experienced situational criticality of takeover situations is in the range of harmless and varies between 1.6 for L2 and 2.4 for the black condition (see figure 2). There is a significant difference with TORs being experienced as less critical in L2 compared to black condition, with L3 and L3+ lying in between (F(3,117)=3.5, p<.05).

Takeover reaction time differs between conditions (F(3, 117)=8.8, p<.001) with L2 having an average reaction time of 2.1 seconds and the other three conditions ranging between 2.5 and 2.7 seconds (figure 2).

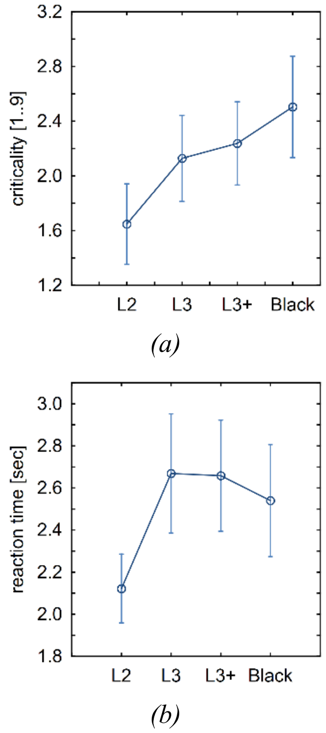



***Figure 2*** *Experienced criticality (a) and reaction time until AD is turned off after a TOR (b).*

## 4. Discussion

Items from the post-drive questionnaire addressing the perception of the driving scene during AD mode show the intended variation of SA with L2 with the highest level, the black condition with the lowest level and L3 and L3+ lying in between. Similarly, in the evaluation of the takeover situations directly after each TOR, the conditions group into three levels with L2 being the least and black the most critical.

However, for items of the post-drive questionnaire that relate to situational aspects perceivable outside the forward view and to perception at a TOR, the differences between conditions are reduced. Now, mainly L2 corresponds to a higher level of SA while the three other conditions become very similar. A similar pattern with a difference between L2 and the three other conditions is found for takeover responses measured by takeover times.

The reported takeover time is a commonly used measure for assessing driver performance at TORs (Yining Cao, Zhou et al., 2021) that reflects the timely component of a takeover response. Still, to fully understand the relation between SA and takeover reaction, indicators are needed (like TOC-rating, Naujoks, Wiedemann, Schömig, Jarosch, & Gold, 2018) that go beyond the reported timely aspect of takeover performance.

One explanation for the two different patterns of differences between conditions could be that during AD mode there were indeed three different levels of SA. In takeover situations, drivers were able to compensate the reduced SA in the black condition so that this condition did no longer differ from the L3 and the L3+ condition. However, the additional effort required during the takeover response is still reflected in the experienced overall criticality of the situation.

## 5. Conclusions

The used post-drive questionnaire was suited to reflect differences in SA. This is especially true for items assessing SA while being in AD mode. Here, results indicate that even the perception of aspects of the environment that do not directly result from enhanced visual attention (like vehicle dynamics, noise) might be positively impacted in L2 compared to L3.

For a final understanding of the relation between the different measures of SA at different points in time (during AD mode, at TOR), the data needs to be analysed in more detail. More indicators like gaze patterns, situational aspects of the takeover situations or a more fine-grained analysis of takeover response will give more detailed insight into the different levels of SA and how they impact drivers' response to TORs.

## 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. Neither the European Commission nor CINEA – in its capacity of Granting Authority – can be made responsible for any use that may be made of the information this document contains.

## References

Cortens, B. (2019). *Impact of Alert Design and Hazard Direction on Driver Behaviour and Understanding in a Simulated Autonomous Vehicle* (Master' thesis, University of Guelph, Guelph, Kanada). Retrieved from https://atrium.lib.uoguelph.ca/xmlui/handle/10214/15935

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. Paper presented at the *Proceedings of the Human Factors Society annual meeting*

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors*, 37(1), 65-84.

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver bach into the loop? Paper presented at the *Human Factors and Ergonomics Society 57th Annual Meeting*, San Diego, CA, USA.

Liang, N., Yang, J., Yu, D., Prakah-Asante, K. O., Curry, R., Blommer, M., . . . Pitts, B. J. (2021). Using eye-tracking to investigate the effects of pre-takeover visual engagement on situation awareness during automated driving. *Accident Analysis & Prevention*, 157, 106143.

Merat, N., & Jamson, A. H. (2009). Is Drivers' Situation Awareness Influenced by a Fully Automated Driving Scenario? Paper presented at the *Human factors, security and safety*.

Moore, K., & Gugerty, L. (2010). Development of a novel measure of situation awareness. The case of eye movement analysis. Paper presented at the *Human Factors and Ergonomics Society Annual Meeting*

Naujoks, F., Wiedemann, K., Schömig, N., Jarosch, O., & Gold, C. (2018). Expert-based controllability assessment of control transitions from automated to manual driving. *MethodsX* 5: 579-592.

Neukum, A., Lübbeke, T., Krüger, H.-P., Mayser, C. & Steinle, J. (2008). ACC-Stop&Go: Fahrerverhalten an funktionalen Systemgrenzen. In M. Maurer & C. Stiller (Eds.), *5. Workshop Fahrerassistenzsysteme - FAS 2008*. pp. 141-150. Karlsruhe: fmrt.

SAE. (2021). J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (Vol. J3016), SAE International/ISO, www.sae.org

Schwindt, S., von Graevenitz, P., & Abendroth, B. (2023). Situational Awareness in the Context of Automated Driving - Adapting the Situational Awareness Global Assessment Technique. Human Factors in Transportation, Vol. 95, 2023, 161–171, https://doi.org/10.54941/ahfe1003802

Sirkin, D., Martelaro, N., Johns, M., & Ju, W. (2017). Toward measurement of situation awareness in autonomous vehicles. Paper presented at the *2017 CHI Conference on Human Factors in Computing Systems*.

Strybel, T. Z., Chiappe, D., Vu, K.-P. L., Miramontes, A., Battiste, H., & Battiste, V. (2016). A comparison of methods for assessing situation awareness in current day and future air traffic management operations: Graphics-based vs text-based online probe systems. *IFAC-PapersOnLine*, 49(19), 31-35.

Taylor, R. M. (2017). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational awareness* (pp. 111-128): Routledge.

van den Beukel, A. P., van der Voort, M. C., & Eger, A. O. (2016). Supporting the changing driver's task: Exploration of interface designs for supervision and intervention in automated driving. *Transportation research part F: Traffic psychology and behaviour, 43*, 279-301. doi:10.1016/j.trf.2016.09.009

van de Merwe, K., van Dijk, H., & Zon, R. (2012). Eye movement as an indicator of situation awareness in a flight simulator experiment. *The International Journal of Aviation Psychology,* 22(1). doi:https://doi.org/10.1080/10508414.2012.635129

Yining Cao, Y., Zhou, F., Molnar, L., Robert, L.P., Tilbury, D.M. & Yang. X. J. Towards Standardized Metrics for Measuring Takeover Performance in Conditionally Automated Driving: A Literature Review, *Proceedings of the 65st Annual Meeting of the Human Factors and Ergonomics Society*, Oct. 4 - 7, 2021 at the Baltimore, MD.

Zhang, T., Yang, J., Liang, N., Pitts, B. J., Prakah-Asante, K. O., Curry, R., . . . Yu, D. (2020). Physiological measurements of situation awareness: a systematic review. *Human Factors*.

# Drive-In Lab: A method for the measurement and rating of OEM in-car infotainment system's distraction potential

Tuomo Kujala[1*], Hilkka Grahn[1], Toni Hautaoja[1], Abhishek Sarkar[1]

[1] Cognitive science, Faculty of Information Technology, University of Jyväskylä, P.O. Box 35, FI-40014 University of Jyväskylä, Jyväskylä, Finland
[*]tuomo.kujala@jyu.fi

**Abstract: This extended abstract introduces the Drive-In Lab at the University of Jyväskylä. The laboratory was constructed in early 2024 for enabling valid and reliable measurement and rating of in-car infotainment system's distraction potential. Any car can be driven into the laboratory and connected to a driving simulation, which enables measurement of in-car tasks' distraction effects in a controlled traffic scenario. It can provide an estimate of an in-car task's effect on crash risk in a car-following scenario as compared to attentive baseline driving. Based on effect size, a 0–3 star rating (large, medium, small, no effect) can be given for each in-car task, and for the whole infotainment system based on the mean effect size across the tasks.**

## 1. Introduction

There is no question that the use of smartphones while driving has a negative impact on driver distraction (e.g., Caird et al., 2014; Guo et al., 2010; Lipovac et al., 2017; Simmons et al., 2016). However, modern in-car infotainment systems have started to rival smartphones in functionality, allowing drivers to perform various tasks beyond driving-related functions.

There is large variability in how the user interfaces (UIs) of these systems function between different car models. Comparative knowledge of the distraction potential of original equipment manufacturer (OEM) infotainment systems is rare. Further, there are no best practice recommendations for modern in-car infotainment UIs. At times, it seems that a leading design principle is user experience instead of minimization of distraction. Partial automation of the driving task or driver monitoring systems do not solve the problems as the driver should still be able to supervise the functioning of the automated driving and drivers might not obey distraction warnings (Lubkowski et al., 2021).

Car manufactures rarely publish their own test results which makes the comparison between OEMs' infotainment systems challenging. Strayer et al. (2015, 2017) benchmarked ten model 2015 cars' infotainment systems on real roads with several metrics. To provide reliable and comparable test results, simulated driving could offer control over undesired confounding variables.

For the reasons above, we are launching a research project for studying the distraction potential of cars' OEM infotainment systems which aims at well-controlled, valid, and reliable testing. We have built a Drive-In Laboratory, where a test vehicle can be driven inside and connected to a driving simulation. Here, we introduce the new laboratory and the associated distraction potential measurement.

## 2. Method

### 2.1 Participants

Based on power analysis, $N = 32$ should be sufficient sample size per car model to ensure statistical significance of medium-sized effects in a paired-sample t-test with power close to .80. To ensure representativeness of the age distribution in the driver population, we will recruit four drivers per age group, following the recommendation by NHTSA (2013); 18–24, 25–39, 40–54, and 55+. Balanced gender distribution is targeted. An eligible participant should have a valid driver's license, normal or corrected vision, and not have any experience with the infotainment system under testing.

Further, participants' visual search speed, headway preference, and brake reaction time (BRT) for an unexpected event are estimated to enable comparison between participant samples between tests. These measures are necessary to avoid

***Fig. 1.*** *Drive-In Lab at the University of Jyväskylä (NB. The images are from the initial construction phases.)*

a situation where the outcome of a test is more dependent on the qualities of a participant sample than on the qualities of the in-car user interface.

### *2.2 Equipment – The Drive-In Lab*

The Drive-In Laboratory is located inside a 6 m x 7.2 m container (see Figure 1). Three BenQ LU960ST projectors project a 5760x1200 front view to a cylindrical 7850x2000 mm screen with a 5000 mm diameter, and one BenQ LH820ST+ projector displays a 1920x1080 back view to a 4000x1950 mm screen. These provide a feeling of immersion and ensure realistic looming effects.

A test car is lifted with a Nussbaum Sprinter Mobil 3000 jack for safety reasons. The car is connected to the driving simulation software with external sensors. Yocto-3D-V2 position sensor is used for reading the steering movements and Moza Sr-p pedals are placed above the car pedals to ensure the same accuracy of longitudinal control between different car models. Steering is automated for the testing, but the position sensor enables analyses of lateral control for other purposes.

Eepsoft Oy provides the lab's professional driving simulator software with extensive data gathering possibilities. Eye-tracking system and behavioural research software are used for data synchronization and transcription. 4k camera records the participant's hand movements and use of controls from the backseat of the car. An Android tablet is used for measuring participant's stationary visual search speed three times with Arrows Task (by Jukic Hrvoje) and for collecting subjective experiences on the tasks and system (e.g., NASA-TLX, SUS, UEQ). For the test results, the driving log data is sufficient, but the other data are valuable for additional research purposes.

### *2.3 Procedure, Measurement and Analysis*

A representative sample of 10 typical in-car tasks is chosen for testing, such as: turn AC to x degrees, tune into radio station y, start navigation to address z. Each task is repeated as many times as the participant can complete for the duration of a 4-minute drive. The participants are not instructed or trained for the use of the systems, to study the worst-case scenario where the driver starts to use the system for the first time while driving. This enables analyses of learning effects and intuitiveness of the UIs.

For the distraction measurement we'll collect distance headway (DHW) data in a car-following scenario. The participant's task is to get efficiently from point A to point B, while keeping safe distance to a lead car. A test starts by a BRT trial where the participant is instructed to keep a minimum DHW to a lead car that the participant thinks is still safe. The lead car keeps its speed constant at 80 km/h, until it suddenly brakes hard with -6 $m/s^2$. Participant's individual BRT for the DHW at the onset of the braking is thereby estimated. After this, the participant completes a 4-min baseline drive of attentive driving, in which the lead car adjusts its speed based on the same algorithm as in the following scenarios with the in-car tasks.

Next, a participant drives the test car surrounded by simulated traffic while conducting in-car tasks. The lead car in front of the participant decelerates and accelerates in seemingly unpredictable manner but for preset durations. The lead car will always decelerate so that the DHW goes under the critical DHW in Eq.1. if the participant does not decelerate, but without crashing.

After the BRT trial, the participant should be aware that it is always possible that the lead car brakes hard. The variable critical thresholds for the DHW are based on this possible hazardous scenario in all situations. To detect distraction, we will measure in-car tasks' effects on DHW adjustments by comparing the headways with in-car tasks to the subjectively preferred headways in the baseline drive without in-car tasks. The assessment is focused on the impact of the in-car task on maintenance of appropriate DHW in relation to a variable critical distance headway $DHW^{(t)}{}_{critical}$ at and below which there is a possibility of a rear-end crash (Kujala & Sarkar, 2024):

$$DHW^{(t)}{}_{critical} = BD^{(t)}{}_F + S^{(t)}{}_F \cdot BRT^{(t)}{}_F - BD^{(t)}{}_L \quad \text{(Eq.1)}$$

where $BD^{(t)}{}_F$ is the braking distance of the following car, $S^{(t)}{}_F$ is the speed of the following car, $BRT^{(t)}{}_F$ is the participant's brake reaction time that is corrected for the situational DHW, and $BD^{(t)}{}_L$ is the braking distance of the lead car, each at time *t*.

With a 10 Hz sampling rate for a 4-min drive, we'll get 2400 data points per drive/in-car task. We'll mark each data point as -1 if there was a possibility of a rear-end collision and 1 if not, based on Eq.1. Then, we'll compare percentages of -1s between the 4-min baseline drive vs. each in-car task drive with paired-samples t-test or Wilcoxon signed rank test, depending on the distributions. Each in-car task gets a 0–3 rating based on the observed effect size (*d* or *r*: large, medium, small, no effect). After a test, the in-car infotainment system will get a combined rating between 0–3 for its distraction potential based on the mean effect size across the 10 in-car tasks.

## 3. Conclusions

The new method can produce estimates of in-car task's effect on crash risk potential in car-following scenarios in controlled laboratory settings. The associated metrics of distraction are based on individual and situationally variable thresholds for distracted driving. The method fulfills all the key requirements for measurement of in-vehicle user interfaces' distraction potential as defined by Kujala and Grahn (2022, see Appendix 1).

Our aim is to offer consumers reliable information on the distraction potential of in-car infotainment systems and to encourage car manufactures to design safer in-car UIs. Further, the testing is expected to support design by revealing which UI solutions are less distracting than others. We wish that in longer term our work has an impact on the development of safety ratings for OEM infotainment systems (cf., Imberger et al., 2020). The first test results are presented at the DDI'24 conference.

## 4. Acknowledgments

This research was supported by the Academy of Finland (Appropriate Uncertainty in Manual and Automated Driving, grant number 343259). Construction of the laboratory was funded by Faculty of Information Technology at the University of Jyväskylä.

**References**

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

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

Imberger, K., Poulter, C., Regan, M., Cunningham, M., & Paine, M. (2020). Considerations for the development of a driver distraction safety rating system for new vehicles. *Journal of Road Safety*, *31*(2). https://doi.org/10.33492/jrs-d-19-00243

Kujala, T. & Grahn, H. (2022). Key requirements and a method for measuring invehicle user interfaces' distraction potential. *The 8th International Conference on Driver Distraction and Inattention*, 60-63. https://ddi2022.org/wp-content/uploads/2022/10/DDI2022-ABSTRACTS.pdf

Kujala, T. & Sarkar, A. (2024). Spare visual capacity and driver inattention in dynamic car-following scenarios. Submitted to *Transportation Research Part F: Traffic Psychology.*

Lubkowski, S. D., Lewis, B. A., Gawron, V. J., Gaydos, T. L., Campbell, K. C., Kirkpatrick, S. A., ... & Cicchino, J. B. (2021). Driver trust in and training for advanced driver assistance systems in Real-World driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *81*, 540-556. https://doi.org/10.1016/j.trf.2021.07.003

National Highway Traffic Safety Administration. (2013). *Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices (NHTSA-2010-0053)*. Washington, D.C.: Department of Transportation – National Highway Traffic Safety Administration.

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

Strayer, D. L., Cooper, J. M., Goethe, R. M., McCarty, M. M., Getty, D., & Biondi, F. (2017). Visual and cognitive demands of using in-vehicle infotainment systems. *AAA Foundation for Traffic Safety*, 1–53.

Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J. R., & Hopman, R. J. (2015). Measuring cognitive distraction in the automobile III: A comparison of ten 2015 in-vehicle information systems. *AAA Foundation for Traffic Safety*, 202–638.

## Appendix A

**Table 1** Analysis on how the proposed method fulfils key requirements for measuring in-vehicle user interfaces' distraction potential (Kujala & Grahn, 2022)

| Requirement | Proposed method | Fulfilled |
|---|---|---|
| 1. Inattention should be evaluated against attentive task performance. | The distraction potential is estimated based on the in-car task's effect on the participant's crash risk as compared to a baseline drive without in-car tasks. | ✓ |
| 2. Inattention should be assessed against the spare attentional capacity available in attentive driving. | The participants can freely glance away from the lead car if needed. These glances are not counted as distraction, regardless of their lengths. However, if distracting, these glances can have an effect on the DHW maintenance. | ✓ |
| 3. Situational variabilities in the spare attentional capacity should be recognized. | The safety-critical DHW to which the participant's situational DHW is compared to, varies by situational relative speed and distance to the lead car (Eq.1). | ✓ |
| 4. Inter-individual differences in the spare attentional capacity should be controlled for. | The safety-critical DHW to which the participant's situational DHW is compared to, varies from participant to participant based on their individual BRTs for an unexpected event (Eq.1). | ✓ |
| 5. Drivers' cognitive processing abilities and limitations should be acknowledged. | The participants can themselves decide on what they believe is a safe DHW. Their individual BRTs are estimated and used for the measurement of distraction. They can glance away from the lead car without being labelled as distracted, unless this has an effect on their DHW maintenance. Their visual search times in a stationary visual search task is estimated and samples consist of drivers from young to old, to keep the samples in balance between tests. The looming effects of the lead car are realistic. | ✓ |
| 6. Evaluation should focus on cognitive processes that are relevant for attentive driving. | Attending the lead car, safe DHW maintenance and longitudinal control of the car, based on anticipating what can be possible in a scenario, are prerequisites for safe driving in the real world. | ✓ |
| 7. Evaluations should be probabilistic to avoid hindsight bias. | The evaluation is based on what could be possible in a situation if the lead car brakes suddenly hard. Crashes are not needed for the distraction potential evaluation. | ✓ |
| 8. There should be a link to real-life crash risk – or to a real-life performance failure probability. | The distraction potential is estimated based on the effect of an in-car task on the participant's crash risk in a baseline drive without in-car tasks. The effect sizes can be linked to an increase in a real-world crash probability in car following. | ✓ |
| 9. Possibility should be more important than probability. | The evaluation is based on what could be possible in a situation if the lead car brakes suddenly hard. The probability of this really happening is not relevant for the test result. What matters for safety in the real world is if a crash is possible when a lead car brakes hard. | ✓ |
| 10. The assessment should be based on the worst-case scenario. | The participants start to learn the in-car tasks while driving without any knowledge of the UI. The testing is based on manual driving in a car-following scenario (no ACC) and on the possible worst-case scenario of a lead car braking suddenly hard. While there could be effects of in-car tasks also in steering performance, steering is automated to be able to better differentiate the distraction effects on participants' longitudinal control performance. | ✓ |

# Empirical Evaluation of Demands Imposed on Drivers by Characteristics of Dynamic Visual Information

Sebastian Gary[1*], Nadja Schömig[1], Dennis Befelein[1], Alexandra Neukum[1], Klaus Perlet[2], Sebastian Hergeth[2]

[1] Wuerzburg Institute for Traffic Sciences (WIVW GmbH), Robert-Bosch-Strasse 4, Veitshöchheim, Germany
[2]BMW Group, Knorrstraße 147, Munich, Germany
[*]gary@wivw.de

**Abstract: The current research investigates attentional demands imposed on the driver by characteristics of dynamic visual information while driving using the example of animations. Both voluntary and involuntary diversion of driver attention towards animated dot clouds were assessed in two separate simulator studies (*n* = 21 each). Animations varied in duration (2s vs. 20s) as well as in attention capturing properties (containing vs. not containing abrupt onsets, looming, luminance contrast and contrast polarity changes). They were presented at random intervals in a car-following task. To assess voluntary driver diverted attention (DDA) induced by the animations, in study 1, participants were instructed to continuously describe the animations. To assess involuntary DDA towards the animations, in study 2, a visual detection response task (DRT) was implemented and participants were instructed to prioritize this task in order to create a top-down goal for attention allocation. Frequency and duration of glances towards the animations were assessed. Reactions to DRT dots presented with a fixed onset after animations were examined to evaluate involuntary diversion of attention. The results are currently under analysis. They will be available and presented at the conference.**

## 1. Introduction

### *1.1 Usage in vehicles*

The use of dynamic visual information such as animations is becoming more widespread in human-machine interface design. Animations can be understood as the "interpolated transition from a visual property value that a 2- or 3-dimensional object has to another property value" (e.g., its extension, position in space, shape, possible texture, colour or transparency; Schilbach, 2014). Animations may serve different purposes: They could help increase understandability of functions, provide feedback about task execution, increase trust and acceptance in a system functionality, or contribute to making a feature more pleasant to use and satisfying.

In vehicles, animations can be used for visualizing both driving-related and non-driving-related information.

### *1.2 Potential effects on driver attention*

Like any other cue, such animations may theoretically divert driver attention, resulting in

- **voluntary Driver Diverted Attention (DDA),** when a driver deliberately (top-down) directs attention to a stimulus,

and / or

- **involuntary DDA**, when a stimulus reflexively diverts attention (bottom-up) away from activities critical for safe driving.

Drivers who deliberately divert attention have some possibility for adapting their behaviour to mitigate interference with driving performance. Such self-regulation may not be possible if a distracting stimulus is compelling and induces involuntary DDA. Accordingly, the mechanisms underlying these categories of engagement might differ and cause distinct interference patterns (Regan et al., 2011).

Even though helpful in categorizing how DDA is initiated, this distinction is rarely considered in the literature (Regan et al, 2011).

### *1.3 Demands imposed on drivers*

The impact of an animation on driver attention will depend on its attentional demand on the one hand and its frequency as well as duration on the other hand. In contrast to the frequency of an animation that may depend on conditions inside and outside of the vehicle, the duration and attentional demands of a given animation can be designed more deliberately.

Previous research indicates a link between driver visual attention and crash risk, for example, risk increases with longer eyes-off-road times (Forster et al., 2024). With respect to voluntary DDA, the duration of an animation may influence the time drivers look at it. We therefore hypothesized that *(H1) animation duration affects driver eyes-off-road times.*

With regard to involuntary DDA, previous research has identified properties of things which could be described as compelling and may capture driver attention (Regan et al., 2011). According to Franconeri and Simons (2003) these are "new objects, objects that move suddenly, and looming objects that are all behaviourally urgent". Another feature that has been consistently shown to capture attention involuntarily is changes in luminance contrast paired with a change in luminance contrast polarity (Franconeri and Simons, 2003). Therefore, we hypothesized that *(H2) animations with abrupt onsets, looming, as well as concurrent changes in luminance contrast and contrast polarity can initiate involuntary DDA*.

Based on previous observations, we further expected that *(H3) effects of animation on driver eyes-off-road times change over time*.

### 1.4 Study goals

This publication presents an empirical evaluation of demands imposed on drivers by characteristics of dynamic visual information. Two driving simulator studies were conducted to investigate voluntary and involuntary DDA, based on different operationalizations of how participants should distribute their attention between the driving and non-driving tasks.

## 2. Method

### 2.1 Study design for distinguishing voluntary and involuntary DDA

Animations containing stimuli that may give rise to voluntary and involuntary DDA were implemented in two studies. Study 1 was designed to assess voluntary DDA towards animations. This was achieved by instructing participants to continuously describe the animations when they were shown. Study 2 was designed to assess involuntary DDA induced by animations. Therefore, high workload was induced by implementing a visual Detection Response task (DRT; ISO, 2016) presented on the front screen of the simulator projection. The driver's task was not to miss any of the presented dots, so that any attention that nevertheless was diverted to a presented animation in this setup can be considered as involuntary. Each study part was conducted with $n = 21$ participants each.

### 2.2 Stimuli characteristics

The exemplary animations investigated in the studies were various geometric objects (e.g. balls, disks, cubes, pyramids, stars) comprising of point clouds presented in the vehicle's central information display (CID, Figure 1).

In each study, a 2 x 2 x 3 factor design was used for varying animation characteristics:

- Factor 1, duration of animation: long (20s) vs. short (2s)
- Factor 2, attention capturing properties: containing vs. not containing abrupt onsets, looming, luminance contrast and contrast polarity changes
- Factor 3, time of measurement: first vs. second vs. third presentation of animation

Combining factor 1 and 2 resulted in four different animation types which were implemented in four separate experimental drives. To investigate habituation effects, each animation type was presented three times within each drive varying in the objects shapes shown.

Table 1 summarizes verbally the animations specified for the four different drives.

***Table 1*** *Specification of the exemplary animation types*

| | | Animation design | |
|---|---|---|---|
| | | NOT containing attention capturing properties | Containing attention capturing properties |
| Animation duration | Short (2s) | One object fades in for 2s, moves slowly around itself for 2s | One object looms in for 2s, after 500ms contrast changes from black to white; during looming object changes its shape |
| | Long (20s) | One object fades in for 2s, moves slowly around itself for 20s | An object looms in for 2s, after 500ms contrast changes from black to white; during looming object changes its shape; In the next 3s, the object moves around itself and becomes smaller<br>4 objects with various shapes are presented with this sequence |

### 2.3 Driving task/layout of experimental drive

Studies were conducted in the WIVW driving simulator using the simulation software SILAB® (see Figure 2). The driving task was a continuous car-following task (slightly adapted from the EGDS protocol described in NHTSA, 2013). Participants were instructed to follow a lead vehicle at a constant speed of 100 km/h with a distance of 50 m as constantly as possible and to maintain good lane keeping. Every 60 – 90 seconds an animation was presented. In total, one drive took five minutes.


**Fig. 1.** Car-following task in the simulator. The red dot represents the visual DRT in study 2. The animations are displayed in the CID. An example of a 2s animation containing attention capturing properties is shown.

### *2.4 DRT specification for study 2*

The DRT in study 2 was implemented for two reasons: Firstly, the DRT should ensure driver attention toward the forward road scene in the chosen experimental setting by emphasizing the relevance not to miss any DRT dot. Secondly, reaction times and missing events were used to operationalize the effects of animations on driver visual and mental workload. Therefore, the DRT was implemented according to ISO specifications (2016; *continuous DRT dots*, every 3-5 seconds).

In addition, the presentation of animations was timed so that an auxiliary DRT dot was presented exactly 800 ms after animation onset to evaluate involuntary attention capture. By analysing these *timed DRT dots* separately from the continuous DRT dots, missing rates and reaction times reflect a distinct measure of involuntary attention shifting caused by the animations. For the participants, the timed DRT dots were not distinguishable from the continuous DRT dots.

### *2.5 Dependent measures*

Dependent variables were driver glance behaviour measured via an eye tracking system by SmartEye® (total number of glances, mean single glance duration, total glance duration towards animations). In addition, reaction time until first glance to animations after onset, reaction times and missing rates for continuous and timed DRT processing were measured as indicators of involuntary attention capture.

### *2.6 Theoretical implications and practical application*

Dynamic visual information such as animations can support the driver, and if implemented judiciously, should not give rise to driver inattention. For example, Birrell and Fowkes (2014) found that an integrated smart driving system providing feedback to the driver via a dynamic, adaptive interface did not induce visual distraction, with monitoring being incorporated into normal driving. Categorizing how DDA may be initiated voluntarily and involuntarily has scarcely been researched but may prove particularly useful when thinking about distraction mitigation in this context.

The current research addresses this hole in the literature by identifying relevant properties of animations and exploring them in empirical studies, thereby adding to the understanding of different forms of inattention.

## 3. Acknowledgments

This research has been funded by the BMW Group. All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations. We thank the editors and reviewers for their valuable time and effort.

## References

Birrell, S. A., & Fowkes, M. (2014). Glance behaviours when using an in-vehicle smart driving aid: A real-world, on-road driving study. Transportation Research Part F: Traffic Psychology and Behaviour, 22, 113–125. doi:10.1016/j.trf.2013.11.003.

Forster, Y., Schoemig, N., Kremer, C., Wiedemann, K., Gary, S., Naujoks, F., ... & Neukum, A. (2024). Attentional warnings caused by driver monitoring systems: how often do they appear and how well are they understood? Accident Analysis & Prevention, 205, 107684.

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. Perception & psychophysics, 65(7), 999-1010.

International Organization 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). Retrieved from https://www.iso.org/standard/59887.html

National Highway Traffic Safety Administration (2013). Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices (Docket No. NHTSA-2010-0053). Washington, DC: U.S. Department of Transportation National Highway Traffic Safety Administration (NHTSA).

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.

Schilbach, J. (2014), An Event-Based Framework for Animations in X3D. In: Proceedings of the 19th International ACM Conference on 3D Web Technologies (Web3D ’14), 89–97, ACM Press, New York, doi:10.1145/2628588.2628595

# The Role of Mental Models on the Effectiveness of Driver Monitoring Systems in Reducing Driver Distraction: Insights from a Driving Simulator Study

Ina Koniakowsky [1,2*], Yannick Forster [1], Josef F. Krems [2], Andreas Keinath[1]

[1] BMW Group, Munich, Germany
[2] Cognitive and Engineering Psychology, Chemnitz University of Technology, Chemnitz, Germany
[*]ina.koniakowsky@bmw.de

**Abstract: Driver monitoring systems (DMS) can detect driver distraction and prompt drivers to keep their eyes on the road. Consumer advocates have introduced a DMS that considers both Long Distraction (LD) and Visual Attention Time Sharing (VATS) to derive driver distraction levels, but empirical evidence on its effectiveness is lacking. Therefore, the current paper presents a driving simulator experiment ($N = 52$) comparing the frequency of LD and VATS with an active and an inactive DMS. The DMS significantly reduced the number of LD and VATS. Although more warnings were generated by VATS, the DMS was less effective in reducing VATS compared to LD. Self-report data revealed that drivers had an incomplete mental model of the DMS, i.e., they did not recognize that VATS triggered warnings. This may explain the smaller effects for VATS and underlines the importance of mental model formation for the effectiveness of a warning system.**

## 1. Introduction

Driver monitoring systems (DMS) have the potential to be a standard safety feature in future vehicles. There is a trend to implement DMS for detecting impaired driving in automated as well as manual driving. In its latest safety protocol, the European New Car Assessment Program (Euro NCAP) considers two types of gaze behavior contributing to distraction: a 3-second shift of the driver's gaze away from the road, i.e., long distraction (LD), or multiple short shifts that accumulate to 10 seconds within the last 30 seconds, i.e. visual attention time sharing (VATS) (EuroNCAP, 2023). The aim of DMS is to mitigate distraction by providing real-time feedback to the driver whenever they are distracted. While the effectiveness of DMS in detecting distraction is well recognized (Dong et al., 2011), the pressing issue of their impact on gaze behavior is still being debated.

Some previous studies on DMS with different algorithms for distraction detection have shown that these can reduce the probability of long gazes (> 2 s) (Atwood et al., 2019; Victor et al., 2018), while others could not demonstrate significant effects of a DMS (Ahlstrom et al., 2013). Regarding Euro NCAP's algorithm, there have been scientific efforts, showing that drivers trigger multiple warnings in both driving-related (Forster et al., in press) and non-driving-related use cases (UC) (Koniakowsky et al., 2023). According to Wogalter (2018), the effectiveness of a warning system depends on the user's understanding, beliefs, and attitudes toward the system. An online survey revealed that only 20 % of drivers believed that DMS would be capable of detecting drivers' gaze (Nees & Liu, 2022).

To bring forth evidence, the present study examines the effectiveness of a DMS, following the Euro NCAP specifications, by comparing the occurrence of distracted behavior with and without a DMS. We hypothesize that the DMS significantly reduces the frequency of LD and VATS. In a multi-method approach, self-reported data on drivers' understanding about DMS help to interpret the findings.

## 2. Method

### 2.1 Participants

Fifty-two participants took part in the experiment ($M_{Age} = 36.33$, 19 female). All held a valid German driver's license and had normal or corrected to normal vision.

### 2.2 Apparatus

The study was conducted in a fix-base driving simulator equipped with a fully functional mock-up including a central display. Five front projectors provided a field of view of 220° and three LED screens displayed the rear view. The driving simulation software SILAB® was used to create a three-lane highway scenario (Krüger et al., 2005). Participants' gaze behavior was recorded using the SmartEye® remote eye-tracking system, including three cameras.

### 2.3 Driver monitoring system

The DMS triggered warnings consisting of both visual and auditory components. These were immediately issued after the driver was classified as being distracted (Fig. 1), either through the detection of LD or VATS (EuroNCAP, 2023). Off-road glances were counted with a latency of 150 ms.

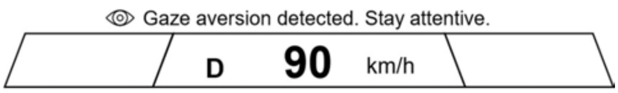



***Fig. 1.*** *Warning above the speed display in the driver's line of sight.*

### 2.4 Design and Procedure

A between-subjects design was employed to examine the effects of a DMS. The DMS was either active generating warnings or inactive not generating warnings. Participants were randomly assigned to one of the groups (DMS active: $n = 26$). During the experiment, participants were asked to repeatedly perform two infotainment-related UCs by touch. The UCs were playing a song from Spotify and making a call. Both required four operating steps.

Participants were neither informed about the DMS nor the occurrence of warnings. They were told that driving safety and compliance with traffic rules always had priority. Upon completion of the drive, participants who received warnings were asked about their understanding of the warnings and potential behavioral changes in a structured interview. Subsequently, they completed a questionnaire on their mental model of the DMS (Beggiato & Krems, 2013).

### *2.5 Analysis*

The number of LD and VATS that occurred from the start of the UCs to its completion served as dependent measures. When the DMS was inactive, VATS and LD were recorded that would have triggered the warnings. The impact of the DMS was inferentially analysed by means of a *t*-test for each measure. Effect sizes are interpreted according to Cohen (1992). Self-report data were analysed based on Mayring (2004). Responses were categorized inductively from the interview material by two independent raters, with high inter-rater reliability ($r = .99$).

## 3. Results

### *3.1 Effect of warnings*

When the DMS was active, drivers received on average 14 warnings while performing ten UCs on the central display ($SD = 6.84$). Warnings were more than twice as often triggered by VATS ($M = 11.04$, $SD = 2.68$) than by LD ($M = 2.96$, $SD = 4.82$). The DMS significantly reduced the number of LD, $t(50) = -1.87$, $p = .033$, $d = 0.52$, which is considered a medium effect. Also the number of VATS was significantly reduced, $t(50) = -1.74$, $p = .044$, $d = 0.48$, which is considered a small effect (Fig. 3).

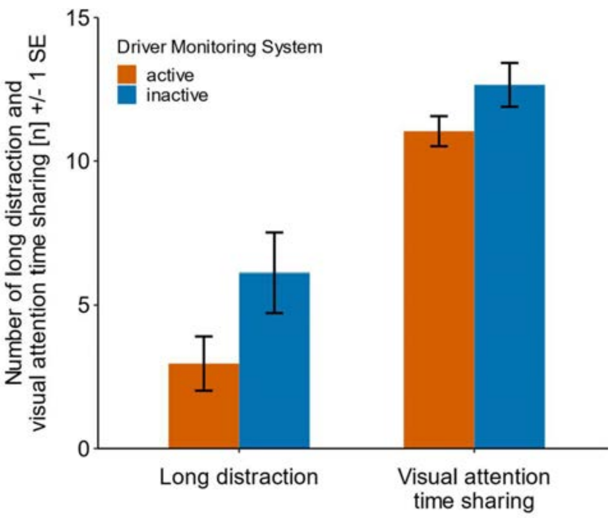



***Fig. 3.*** *Number of LD and VATS by DMS status.*

### *3.2 Understanding of warnings*

Most participants in the DMS condition correctly understood that the warnings were solely triggered by their gaze behavior (69 %). Some believed that the warnings were also related to their driving behavior, like lane departure (19 %). Twelve percent of participants did not understand that the warnings were triggered by their gaze behavior.

In a survey of the mental model (Fig. 2), participants showed a high level of agreement on the statement that long glances triggered warnings ($M = 5.50$, $SD = 0.86$). In contrast,

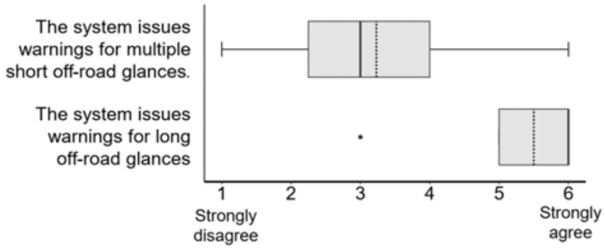



***Fig. 2.*** *Drivers' mental model of the DMS.*

participants were unsure whether multiple short glances caused warnings, which is reflected in the wide range of responses ($M = 3.23$, $SD = 1.27$).

### *3.3 Self-reported behavioral changes*

About two-thirds of participants in the DMS condition indicated they had changed their gaze and/or operating behavior because of the warnings (69 %). Behavioral changes were described as looking at the road more often and for longer periods ($n = 16$), or blind operation ($n = 4$). Further, participants reported that the warnings caused them to interrupt the UC execution more frequently ($n = 7$) or slowed down their execution ($n = 4$). A third reported no behavioral changes and ignored the warnings (31 %).

## 4. Discussion

In summary, there were significant differences in the number of LD and VATS as a function of DMS activation, i.e., the DMS mitigated distracting behavior. However, the effects differed in their magnitude, meaning that the effect of DMS was greater for LD than for VATS. The DMS reduced the occurrence of LD by 52 %, whereas for VATS the reduction was only 13 %. This raises the question, why warnings were more effective for LD than VATS even though they were more frequently triggered by VATS.

Although most drivers were able to link the warnings to their gaze behavior, data further revealed that drivers' mental model of DMS was that warnings were triggered solely by long glances. The participants were not aware that VATS would also elicit a warning, although most warnings were elicited by VATS.

The correct understanding of the warning's intention is a prerequisite for the effectiveness of a warning system (Wogalter, 2018). Only if it is clear which behavior should be avoided, a driver can change accordingly. The incomplete mental model could therefore explain smaller effects for VATS than LD. Similar problems were described in a previous study, demonstrating that participants were unable to build a correct mental model of VATS warnings, even though the algorithm was explained to them (Forster et al., in press). This suggests that VATS warnings may ultimately be too complex to develop a correct mental model. Further research should address the question of how to ensure the efficacy of DMS, to undearstand and avoid VATS warnings.

## 5. Conclusions

This work presents findings that a DMS, using the EuroNCAP algorithm, can mitigate driver distraction. The results highlight the importance of keeping the complexity of an algorithm as low as possible so that drivers can build a correct mental model, as this can be a factor that reduces the effectiveness of the system.

## References


Ahlstrom, C., Kircher, K., & Kircher, A. (2013). A Gaze-Based Driver Distraction Warning System and Its Effect on Visual Behavior. *IEEE Transactions on Intelligent Transportation Systems*, *14*(2), 965–973. https://doi.org/10.1109/TITS.2013.2247759

Atwood, J. R., Guo, F., & Blanco, M. (2019). Evaluate driver response to active warning system in level-2 automated vehicles. *Accident Analysis & Prevention*, *128*, 132–138. https://doi.org/10.1016/j.aap.2019.03.010

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. https://doi.org/10.1016/j.trf.2012.12.006

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Dong, Y., Hu, Z., Uchimura, K., & Murayama, N. (2011). Driver Inattention Monitoring System for Intelligent Vehicles: A Review. *IEEE Transactions on Intelligent Transportation Systems*, *12*(2), 596–614. https://doi.org/10.1109/TITS.2010.2092770

EuroNCAP. (2023). *Assessment Protocol- Safety Assist—Safe Driving. Implementation 2023. V. 10.0.1* (Issue (Version 10.1.1; p. 35)).

Forster, Y., Schömig, N., Kremer, C., Wiedemann, K., Gary, S., Naujoks, F., Keinath, A., & Neukum, A. (in press). Attentional Warnings caused by Driver Monitoring Systems: How often do they appear and how well are they understood? *Accident Analysis and Prevention*. In press.

Koniakowsky, I. M., Forster, Y., Naujoks, F., Krems, J. F., & Keinath, A. (2023). How Do Automation Modes Influence the Frequency of Advanced Driver Distraction Warnings? A Simulator Study. *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 111–114. https://doi.org/10.1145/3581961.3609888

Krüger, H.-P., Grein, M., Kaussner, A., Mark, C., & others. (2005). SILAB-A task-oriented driving simulation. *Driving Simulation Conference*.

Mayring, P. (2004). *Qualitative content analysis* (Vol. 1). Frankfurt: SAGE.

Nees, M. A., & Liu, C. (2022). Mental models of driver monitoring systems: Perceptions of monitoring capabilities in an online U.S-based sample. *Transportation Research Part F: Traffic Psychology and Behaviour*, *91*, 484–498. https://doi.org/10.1016/j.trf.2022.10.012

Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., Aust, & Mikael Ljung. (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

Wogalter, M. S. (2018). Communication-human information processing (C-HIP) model. In *Forensic Human Factors and Ergonomics* (pp. 33–49). CRC Press.

# From Concept to Action - Measuring General and Applied Mental Models in the Context of Automated Driving

Sarah Schwindt-Drews [1*], Bettina Abendroth [1]

[1] Technical University of Darmstadt, Department of Mechanical Engineering, Institute of Ergonomics and Human Factors, Otto-Berndt-Straße 2, D-64289 Darmstadt, Germany
[*]sarah_selina.schwindt@tu-darmstadt.de

**Abstract: This paper presents a research concept for comparing general and applied mental models in automated driving, with a focus on the transition between automation levels. The research concept measures general and applied mental models, gaze movement, and driving performance within a driving simulator. It aims to correlate different mental models with driving performance, to identify how mental models should be characterized for safe interaction, and to provide insights for developing effective training concepts to improve user interaction with automated systems.**

## 1. Introduction

### *1.1 Mode Confusion and Out-of-Loop Problem*

The ongoing automation of vehicles provides drivers with increasing comfort, but also presents significant challenges (SAE International, 2021). Conditionally Automated Driving (CAD, Level 3) (SAE International, 2021) takes over both longitudinal and lateral control of the vehicle and is capable of recognizing system limitations and prompting the driver to take over driving tasks. While CAD allows the driver to disengage from the driving task and focus on activities, such as reading or texting, it also requires the driver to immediately return attention to the driving task and assume full control of the vehicle in the event of a Takeover Request (TOR). Furthermore, CAD is only available under certain conditions, so that in other cases only partially automated driving (PAD, Level 2) or even no automation can be activated. PAD (SAE International, 2021) also provides longitudinal and lateral control of the vehicle, with the difference that the driver is responsible for monitoring the system and environment. Transitions between these levels not only create out-of-loop problems for the driver, but also mode confusion (Kurpiers et al., 2020). Thus, it is essential for drivers to perceive and comprehend relevant information to ensure safe operation of automated vehicles. Therefore, individuals require a suitable mental model of the autonomous vehicle (Endsley, 2017).

### *1.2 Mental Models in the Automated Driving Context*

Mental models are cognitive representations of an external reality and necessary for real-world orientation (Johnson - Laird, 1980). They enable the categorization of perceived information, and support the comprehension of goals, processes, as well as performance and limitations of systems (Seppelt & Victor, 2020). They evolve with increasing experience and are continuously adjusted (Beggiato & Krems, 2013).

Mental models can be categorized into three types (Fig. 1): conceptual, general, and applied mental models. In the context of automated driving, these can be explained as follows. *Conceptual mental models* are precise and comprehensive representations (Norman, 1983) of vehicles, including the interaction of all sensors and actuators installed. *General mental models* comprise the theoretically and practically acquired knowledge about the goals, processes, structures, and limitations of the vehicles (Seppelt & Victor, 2020) and reflect the driver's understanding of their functions and limitations. The driver's general mental model directs the allocation of attention and thus influences the perception of information, which in turn activates the *applied mental model* (Seppelt & Victor, 2020). The applied mental model is represented by the situation awareness, i.e., the perception, understanding, and projection of a situation, and is reflected in the driver's behavior. However, it is possible that the general mental model and the applied mental model may not align.

### *1.3 Measurement of Mental Models*

Several qualitative and quantitative methods exist for measuring mental models, each with specific advantages and limitations (Beggiato, 2015; Bellet et al., 2009; Kearney & Kaplan, 1997; Richardson et al., 2019; Tergan, 1986). While qualitative methods better represent the development process and individual differences in mental models, quantitative methods provide statistical comparability.

In the field of automated driving, research focuses on investigating the evolution of general mental models with increasing practical experience, depending on the accuracy of the initial vehicle description (Beggiato & Krems, 2013; Beggiato et al., 2015; Blömacher et al., 2018, 2020; Forster et al., 2019; Gaspar et al., 2021). Mental models were usually measured objectively through pre- and post-drive questionnaires that cover some driving functions, limitations, and parts of the interaction concept. However, there has been a lack of comparative analysis between subjectively recorded general mental models and applied mental models, as well as

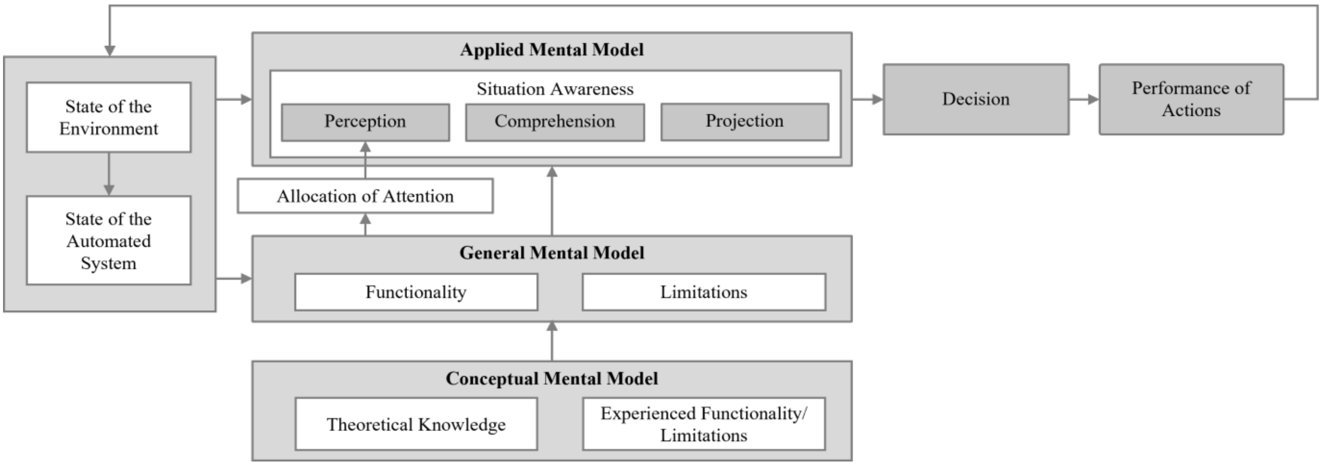



***Fig. 1***. *Conceptual, general and applied mental models in the context of automated driving (based on: Endsley, 2015, 2017; Norman, 1983; Seppelt & Victor, 2020)*

the resulting driving performance. Additionally, mental models have primarily been described for a single level of automation or driver assistance system, rather than the entire automated driving system.

## 2. Research Objective

Given the research gap described above, this paper presents a research concept that enables the measurement and comparative evaluation of the general and applied mental model of the automated driving system and the resulting behavior represented by gaze movement and driving performance. In particular, the change between the automation levels is addressed. The resulting data will provide insights on how a mental model should be characterized to ensure safe interaction with the automated driving system. Based on this, the results will enable the development of a training concept for the education of future users.

## 3. Study design for measuring general and applied mental models

### *3.1 Dependent, Independent, and Confounding Variables and Measurement Methodologies*

The dependent variables to be measured include situation awareness resulting from the applied mental model, as well as the driver's behavior in terms of gaze movement and driving performance (Zhang et al., 2021). Situation awareness is objectively assessed using the Situational Awareness Global Assessment Technique (SAGAT) (Endsley, 1988). The driver's gaze movement is measured using eye tracking (Forster et al., 2019). In order to quantify driving performance, reaction times, time to collision, braking and acceleration behavior, as well as steering behavior are extracted from the driving data (Müller, 2020).

The initial general mental model as measured using the Structural Laying Technique (Scheele & Groeben, 2010) serves as the independent variable between participants. The level of automation activated (Level 0, Level 2, or Level 3) serves as the independent variable that varies within a participant.

Confounding variables include socio-demographic characteristics, driving experience, and experience with automated driving functions, and are collected through questionnaires. Furthermore, reaction time is measured using a stimulus-response test (Matheus & Svegliato, 2013) and motion sickness is assessed pre- and post-driving using the Simulator Sickness Questionnaire (Kennedy et al., 2009).

### *3.2 Experimental Environment*

To create a safe testing environment and capture the applied mental model represented by situation awareness using SAGAT, a fixed-base driving simulator with 360° simulation is selected. The SILAB simulation software is used to conduct a continuous drive with an automated driving system in which the participants experience the transition between automation level 0, 2, and 3 (Fig. 2). Reasons for the level transitions are system limitations such as road section category, roadworks or inappropriate maximum speed limitation.

### *3.3 Procedure*

After informing the participants about the objectives and procedures of the experiment, socio-demographic characteristics, driving experience, experience with automated driving functions, and individual reaction times are recorded. The general mental model is then captured using the Structural Laying Technique. Following this, the participants are given a short introduction on how to operate the vehicle, including the activation and deactivation of the different levels of automation. Simulator sickness is then assessed before participants are equipped with the eye-tracking device and instructed to enter the driving simulator. After a ten-minute familiarization phase, the continuous automated drive begins. During the drive, participants are required to play games on their smartphones whenever it is allowed to engage in a non-driving-related task. Shortly before each level transition, the simulation is paused and the applied mental model, represented by the situation awareness, is assessed using SAGAT. Gaze movement and driving performance are recorded throughout the whole drive. At the end, the participants' simulator sickness status is checked in order to exclude participants significantly affected by simulator sickness.

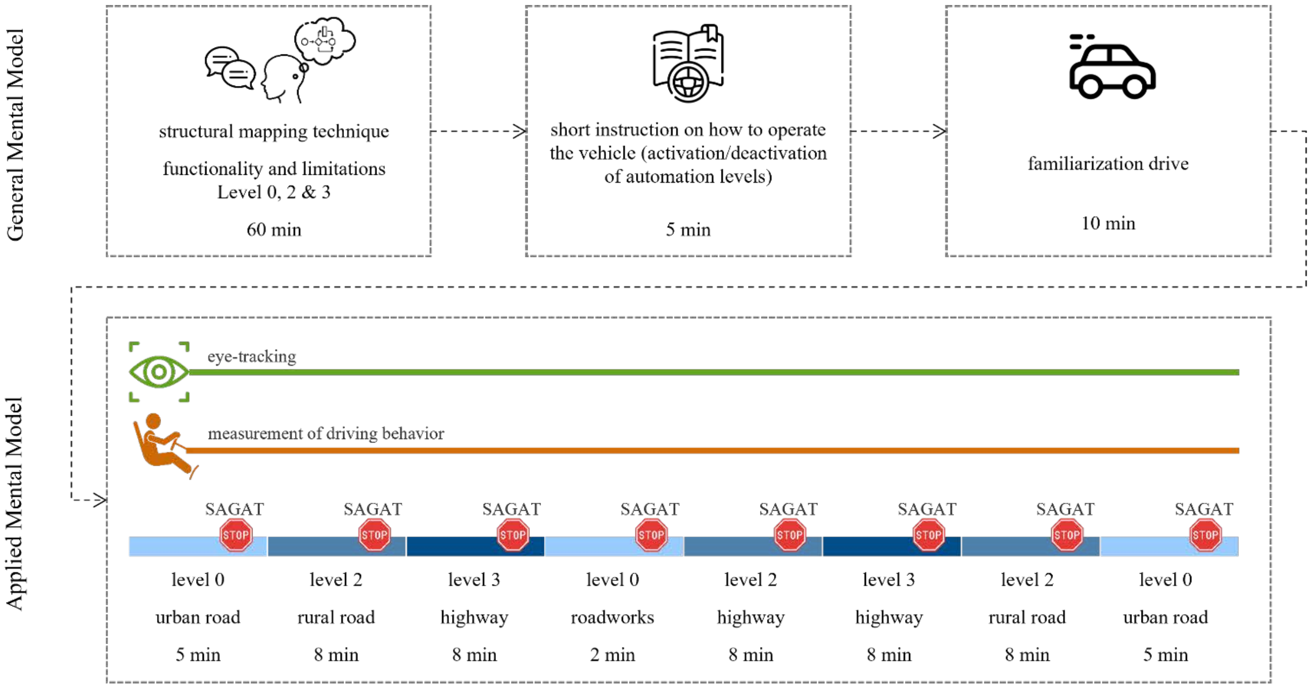



***Fig. 2.*** *Study design for measuring general and applied mental models*

## 4. Advantages and Limitations

Although the validity of the proposed research concept has not yet been tested, this theoretically sound approach provides a way to collect and compare general and applied mental models for automated driving systems. The results are limited by the reduction in realism due to the implementation within a driving simulator. However, an objective measurement of situational awareness using SAGAT is only feasible within a simulation environment (Endsley, 1988). Furthermore, since the research concept provides a relative comparison of mental models, the results can be used without restrictions.

## 5. Conclusions and Future Work

This paper presents a research concept for measuring the general mental model, applied mental model, as well as the resulting gaze movement and driving performance while driving with an automated vehicle, with special focus on the transition between the automation levels. The collected data will provide insights on how a general mental model should be characterized in order to ensure safe interaction with automated vehicles. This will serve as a baseline for developing training concepts to support future drivers. The proposed research concept will be validated through user studies in the next step.

## References

Beggiato, M. (2015). *Changes in motivational and higher level cognitive processes when interacting with in-vehicle automation* [Dissertation]. TU Chemnitz, Chemnitz.

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. https://doi.org/10.1016/j.trf.2012.12.006

Beggiato, M., Pereira, M., Petzoldt, T., & Krems, J. F. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *35*(1–3), 75–84. https://doi.org/10.1016/j.trf.2015.10.005

Bellet, T., Bailly-Asuni, B., Mayenobe, P., & Banet, A. (2009). A theoretical and methodological framework for studying and modelling drivers' mental representations. *Safety Science*, *47*(9), 1205–1221. https://doi.org/10.1016/j.ssci.2009.03.014

Blömacher, K., Nöcker, G., & Huff, M. (2018). The role of system description for conditionally automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, *54*, 159–170. https://doi.org/10.1016/j.trf.2018.01.010

Blömacher, K., Nöcker, G., & Huff, M. (2020). The evolution of mental models in relation to initial information while driving automated. *Transportation Research Part F: Traffic Psychology and Behaviour*, *68*(1), 198–217. https://doi.org/10.1016/j.trf.2019.11.003

Endsley, M. (1988). Situation Awareness Global Assessment Technique (SAGAT). *Proceedings of the IEEE 1988 National Aerospace and Electronics Conference (S. 789–795).*

Endsley, M. (2015). Situation Awareness Misconceptions and Misunderstandings. *JOURNAL of COGNITIVE ENGINEERING and DECISION MAKING*, *9*(1), 4–32. https://doi.org/10.1177/1555343415572631

Endsley, M. (2017). From Here to Autonomy. *HUMAN FACTORS*, *59*(1), 5–27. https://doi.org/10.1177/0018720816681350

Forster, Y., Hergeth, S., Naujoks, F., Beggiato, M., Krems, J. F., & Keinath, A. (2019). Learning and Development of Mental Models during Interactions with

Driving Automation: A Simulator Study. In *Proceedings of the 10th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design: driving assessment 2019* (pp. 398–404). University of Iowa. https://doi.org/10.17077/drivingassessment.1724
Gaspar, J. G., Carney, C., Shull, E., & Horrey, W. J [W. J.] (2021). Mapping drivers' mental models of adaptive cruise control to performance. *Transportation Research. Part F, Traffic Psychology and Behaviour*, *81*, 622–638. https://doi.org/10.1016/j.trf.2021.07.012
Johnson-Laird, P. N. (1980). Mental Models in Cognitive Science. *Cognitive Science*, *4*(1), 71–115. https://doi.org/10.1207/s15516709cog0401_4
Kearney, A. R., & Kaplan, S. (1997). Toward a Methodology for the Measurement of Knowledge Structures of Ordinary People. *Environment and Behavior*(Vol. 29 No.5), 579–617.
Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (2009). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, *3*(3), 203–220. https://doi.org/10.1207/s15327108ijap0303_3
Kurpiers, C., Biebl, B., Mejia Hernandez, J., & Raisch, F. (2020). Mode Awareness and Automated Driving—What Is It and How Can It Be Measured? *INFORMATION*, *11*(5), 277. https://doi.org/10.3390/info11050277
Matheus, C. C., & Svegliato, J. (2013). Stimulus-response tests: An applied demonstration: Demonstration paper. In I. Staff (Ed.), *2013 IEEE Seventh International Conference on Research Challenges in Information Science* (pp. 1–2). IEEE. https://doi.org/10.1109/RCIS.2013.6577740
Müller, A. L. (2020). *Auswirkungen von naturalistischen fahrfremden Tätigkeiten während hochautomatisierter Fahrt* [Dissertation]. TU Darmstadt, Darmstadt. https://doi.org/10.25534/TUPRINTS-00011342
Norman, D. A. (1983). Some Observations on Mental Models. In D. Gentner & A. L. Stevens (Eds.), *Mental Models.* Psychology Press.
Richardson, N. T., Sprung, A., & Michel, B. (2019). Entwicklung und Validierung eines Fragebogens zur Erfassung des mentalen Modells für das hochautomatisierte Fahren. *Zeitschrift Für Arbeitswissenschaft*, *73*(3), 312–323. https://doi.org/10.1007/s41449-019-00157-y
SAE International (2021). *Surface Vehicle Recommended Practice:* Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J3016_202104).
Scheele, B., & Groeben, N. (2010). Dialog-Konsens-Methoden. In G. Mey & K. Mruck (Eds.), *Handbuch qualitative Forschung in der Psychologie* (1. Auflage, pp. 506–523). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92052-8_36
Seppelt, B., & Victor, V. W. (2020). Driver's Mental Model of Vehicle Automation. In D. L. Fisher, W. J. Horrey, J. D. Lee, & M. A. Regan (Eds.), *Handbook of human factors for automated, connected, and intelligent vehicles* (pp. 55–66). CRC Press.
Tergan, S. (1986). *Modelle der Wissensrepräsentation als Grundlage Qualitativer Wissensdiagnostik*. *Beiträge Zur Psychologischen Forschung Ser: v.7*. VS Verlag fur Sozialwissenschaften GmbH.
Zhang, T., Kaber, D., & Zahabi, M. (2021). Using situation awareness measures to characterize mental models in an inductive reasoning task. *Theoretical Issues in Ergonomics Science*, 1–24. https://doi.org/10.1080/1463922X.2021.1885083

# Collecting key performance indicators for distracted driving in Europe: from Baseline to Trendline

Sofie Boets[1*], Agnieszka Stelling[2]

[1] Vias institute, Haachtsesteenweg 1405, 1130 Brussels, Belgium
[2] SWOV Institute for Road Safety Research, Henri Faasdreef 312 2492 JP The Hague, The Netherlands
[*]sofie.boets@vias.be

**Abstract: Driver distraction is a significant risk factor in traffic. Tasks that require looking away from the road and performing manual actions have the greatest impact on driving behaviour and crash risk. The European Commission has defined a key performance indicator (KPI), defined as 'the percentage of drivers not using a handheld mobile device', to monitor the distraction problem. Within the Baseline project minimum methodological guidelines for harmonized data collection and reporting of this KPI were determined. In 2022, fifteen Member States (MS) provided results for the KPI Distraction. Data were mostly collected through roadside measurements by observers; some MS used cameras. The national mean for car, light goods vehicle (LVG) and bus drivers together ranged from 90.6% to 98.3%. Separate indicators were available by road type, week period, vehicle type, age and gender. Distraction was significantly more prevalent in LGV drivers. The methodological requirements generally proved to be feasible, but comparability was not completely reached due to national differences in sampling and weighting. Based on the lessons learnt in Baseline, the minimum methodological guidelines for the KPI Distraction were updated as a part of the follow-up Trendline project. Trendline is a 3-year project, started on 15th October 2022, which brings together twenty-nine European countries for 1) harmonized data collection and reporting of road safety KPIs including distraction and for 2) using them within road safety policies. Data collection within Trendline will take place from 2023 to 2024 and the results will be reported in 2025.**

## 1. Introduction

In 2019, the European Commission (EC) proposed a new approach for the road safety policy for 2021-2030 emphasising the need to use a range of new road safety related Key Performance Indicators (KPIs) (European Commission, 2019). Eight KPIs were defined, referring to main road safety challenges to be tackled, namely: (1) infrastructure safety, (2) vehicle safety, (3) emergency response, road user behaviour with regard to (4) speed, (5) alcohol, (6) distraction, and the use of (7) restraint systems and (8) helmets. The aim of using these KPIs is to monitor the trends in factors that contribute to reaching the EC targets for road safety: moving close to zero fatalities and serious injuries in road transport by 2050 ("Vision Zero") with interim targets of reducing the number of road deaths and seriously injured by 50% between 2020 and 2030.

Driver distraction is included among the proposed KPIs as this is a significant risk factor in traffic (European Commission, 2021). Tasks that require looking away from the road and performing manual actions have the greatest impact on driving behaviour and crash risk. In large scale naturalistic driving research, the crash risk increased by 12.2 times when using a mobile phone in the hand for dialling and 6.1 times when texting (Dingus et al., 2016). Because of the increased use of mobile devices (mainly smartphones) and the widespread use of texting applications, the EC proposed to use the "percentage of drivers not using a handheld mobile device" as proxy for assessing the driver distraction problem (European Commission, 2019).

Within the Baseline project (https://baseline.vias.be), co-funded by the EC, guidelines were developed for the Member States (MS) for the harmonized data collection and reporting of the KPIs, including distraction (Boets et al., 2021).

The aim of the Baseline project was to support MS in providing KPIs, including for driver distraction, in a harmonised way. Detailed methodological guidelines were developed, including the minimum requirements to provide the KPIs as well as optional recommendations. Besides the provision of KPIs by MS and the reporting on the European benchmarking of the KPIs, another aim of Baseline was to evaluate the feasibility and limitations of collecting comparable KPIs across Europe.

## 2. Method

Minimum methodological requirements for the KPI Distraction were set by the EC (2019) and within Baseline (Boets, 2021). The EC (2019) determined the KPI definition (percentage drivers not using a handheld mobile device), method (direct observation by trained observers or from moving vehicles, or alternatives like automatic detection), road types (urban, rural roads, motorways) and vehicle types (cars, light goods vehicles (LGV), buses/coaches) to include, moment of the observations (daylight), and the location sampling procedure (random).

Within Baseline the minimum requirements were further elaborated and operationalised (Boets et al., 2021). The minimum KPIs were set to be for the three vehicle types together and for weekdays (Monday-Friday): 1) the weighted national mean and 2) indicators per road stratum. Only drivers in movement (not stopped) should be included. The locations should allow a good view, inconspicuous observations and be safe. Sample size requirements were: 10 different locations per road type (at least 30 minutes' observation per location), 2,000 drivers in total and 500 drivers per road type. The fieldwork should include a mix of daytime hours on week-and weekend days, balanced over the road types. Fieldwork should not be done during holidays or heavy winter periods. And finally, minimum weighting requirements were determined aiming at providing representative indicators for the country (Silverans & Boets, 2021). Furthermore, recommendations were included: use week periods (weekday and weekend day) as sampling strata and provide separate KPIs (prerequisite: minimum 10 locations per period, 2 locations per period x road stratum, 500 drivers per period), provide KPIs per vehicle type (if min. 500 drivers), boost the sample sizes for more accurate and detailed (crossed strata) estimates, collect data on driver sex and age category and use available national traffic volume estimates (per stratum) to allow proportional data sampling or to weigh the data properly.

## 3. Results

Fifteen European MS collected KPI Distraction data during fieldwork between 2019 and 2022: Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Finland, Germany, Greece, Latvia, Lithuania, Malta, Poland, Portugal, Spain and Sweden (Boets, 2023). Thirteen MS used observers along the road and two MS used camera images (Finland, Lithuania).

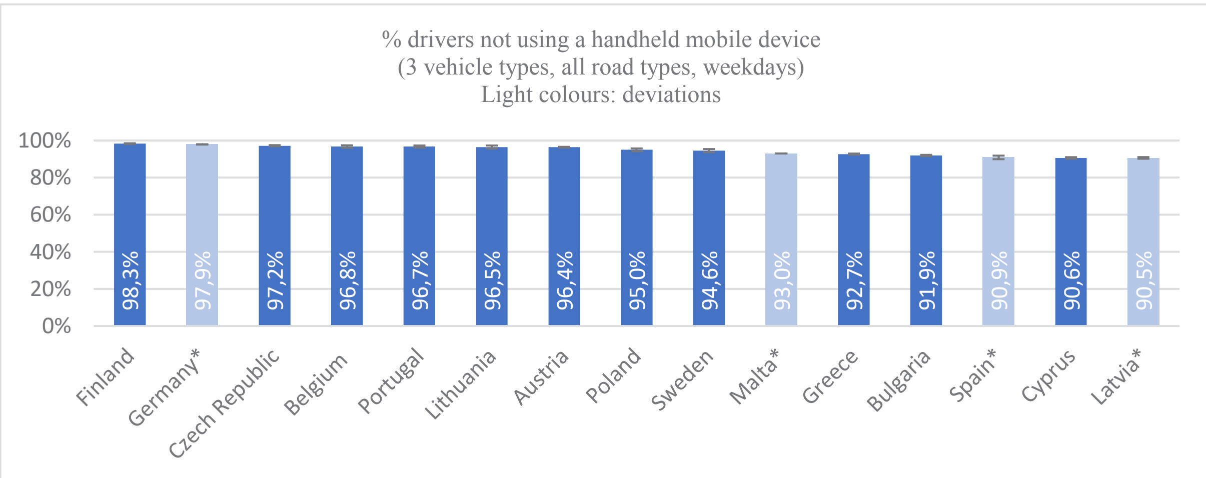
The Baseline experience indicated that the minimum methodological requirements for the KPI Distraction were feasible for most MS. Collecting comparable KPIs for this KPI proved to be possible to a certain extent but not completely due to national differences in sampling and weighting. Furthermore, cameras proved to be possible data collection tools if certain challenges are considered.

Figure 1 shows the national mean KPIs Distraction with 95%-confidence intervals for the three vehicle types (cars, LGV, buses/coaches) and road types (urban, rural roads, motorways) together on weekdays. Light colours indicate deviations from the methodological requirements which are shortly explained underneath the figure. The concerned KPIs cannot reliably be compared with the other presented KPIs. The results indicate that overall more than 90% of the drivers do not use a handheld mobile device while driving. The actual percentages range between 90.6% in Cyprus and 98.3% in Finland.

Besides the minimum KPIs, many MS provided additional indicators. An interesting general pattern among MS (except in Greece) was found with regard to the optional KPIs per vehicle type, namely that LGV drivers more often use a handheld mobile device while driving than car and bus drivers (see Figure 2). Another general pattern concerned the age categories, although only available for three MS, with clearly less handheld mobile device use in 65-plus drivers compared to younger drivers. With regard to the road types and week periods, different patterns were found in the MS, without a common line.

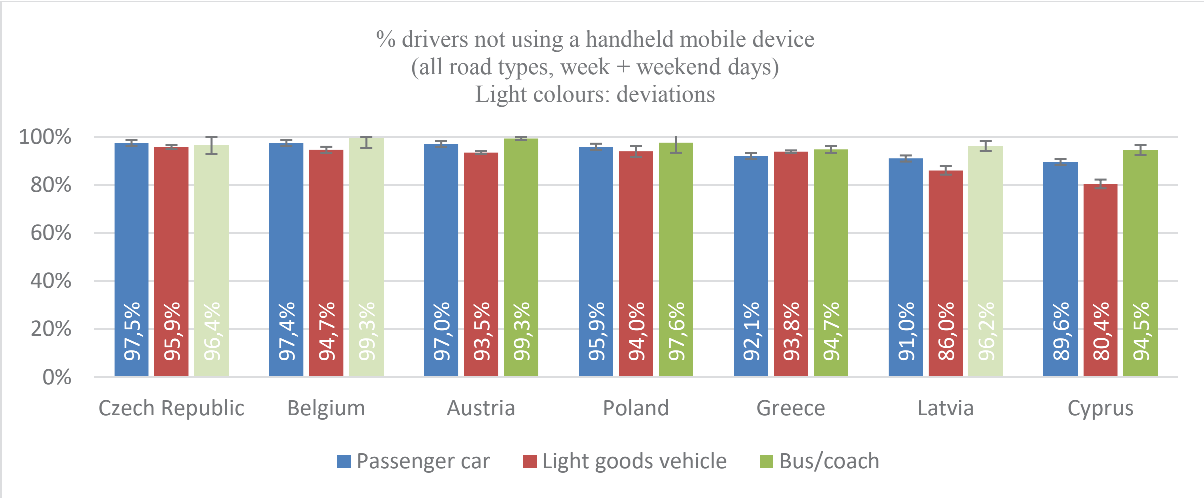
## 4. Discussion and conclusions

Based on the lessons learnt in Baseline, the minimum methodological guidelines for the KPI Distraction were updated as a part of the follow-up Trendline project (https://trendlineproject.eu/). The changes in the methodology will be presented at the conference. Trendline started on 15th October 2022 as a 3-year project which brings together twenty-nine European countries for 1) harmonized data collection and reporting of road safety KPIs including distraction and for 2) using them within road safety policies. Data collection within Trendline will take place in 2023 and 2024 and the results will be reported in 2025. The outcomes of the Baseline and Trendline projects are used to set future European road safety targets and goals based on the KPIs




*Malta, Latvia: no motorways in road network. *Latvia: week + weekend days. *Germany: only passenger cars. * Spain: broader KPI: % having in the hand or operating with the hand a mobile phone or other electronic devices, whether mobile or on-board. * Spain: 4 road types with expressways. *Austria, Greece, Cyprus: % not using a handheld mobile 'phone'. *Finland, Lithuania: based on analysis of camera images; other MS: based on roadside observations by trained observers.

***Fig. 1 -** National Baseline KPIs Distraction*




*Latvia= no motorways. *Germany: only passenger cars. *Austria, Greece, Cyprus: % not using a handheld mobile phone. *Light coloured: deviating method (sample size).

***Fig. 2 -** Baseline KPIs Distraction by vehicle type*


## 5. Acknowledgments

The Baseline and Trendline projects are 50% subsidised by the European Commission and 50% by the national government of the participating Member States.


## References


Boets, S. (2023). Baseline report on the KPI Distraction. Baseline project, Brussels: Vias institute. https://baseline.vias.be/storage/minisites/baseline-kpi-distraction.pdf

Boets, S., Schumacher, M., Stelling, A., Jankowska-Karpa, D. & Pavlou, D. (2021). Methodological guidelines – KPI Distraction. Baseline project, Brussels: Vias institute. https://baseline.vias.be/en/publications/methodological-guidelines-kpi/

European Commission (2019) Staff working document EU road Safety Policy Framework 2021-2030 – Next steps towards "Vision Zero". SWD (2019) 283 final, 32p. Brussels, European Commission. https://transport.ec.europa.eu/system/files/2021-10/SWD2190283.pdf

European Commission (2022) Road safety thematic report – Driver distraction. European Road Safety Observatory. Brussels, European Commission, Directorate General for Transport. https://road-safety.transport.ec.europa.eu/system/files/2022-04/Road_Safety_Thematic_Report_Driver_Distraction_2022.pdf

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., 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 of the United States of America (PNAS), 113(10), 2636-2641. https://doi.org/10.1073/pnas.1513271113

Silverans, P., & Boets, S. (2021). Considerations for sampling weights. Baseline project, Brussels: Vias institute. https://baseline.vias.be/storage/minisites/baseline-considerations-for-sampling-weights.pdf

# Cell Phone Conversations While Driving A Heavy Vehicle: Risk as a Function of Event Type

Jeffrey S. Hickman[1*], Raina Shah[1], Susan Soccolich[2], Rebecca Hammond[2], and Jason Jupe[3]

[1] Applied Safety and Ergonomics, A Rimkus Company, 3015 Miller Rd, Ann Arbor, MI, U.S.A.
[2] Virginia Tech Transportation Institute, 3500 Transportation Research Plaza, Blacksburg, VA, U.S.A.
[3] Rimkus, 8200 Cameron Rd, Suite C-140, Austin, TX, U.S.A.
[*]jhickman@appliedsafety.com

**Abstract: This study evaluated the risk of a cell phone conversation by event type (e.g., sideswipe, etc.) in a sample of commercial motor vehicle (CMV) drivers. Prior naturalistic driving studies have combined event types in their analyses; thus, it is possible that cell phone conversations while driving have different odds ratio estimates based on event type. This has important safety, policy, and regulatory implications as naturalistic driving studies have found that hands-free and handheld cell phone conversations while driving did not increase the odds of a safety-critical event (SCE, crash, near crash, etc.). However, more prevalent event types could bias these results. The study is a re-analysis of existing naturalistic truck driving data from Olson et al. (2009) and Hammond et al (2021). The re-analysis of the SCE data were stratified by different event types. Overall, handheld and hands-free cell phone conversations while driving a CMV did not increase the odds of a SCE compared to no cell phone conversation except for handheld conversations, which significantly increased the odds of involvement in a road departure. Hands-free cell phone conversations, regardless of event type, were largely found to have a protective effect (i.e., decreased the odds of involvement in a SCE). These results support existing regulations (federal and state) for CMV drivers which allow hands-free cell phone conversations while driving.**

## 1. Introduction

In the U.S., there were 415,000 large truck crashes in 2020, of which 101,000 were injury crashes and 4,444 were fatal crashes. Approximately 10 percent of large truck crashes involving a serious injury or fatality were due to truck driver distraction (FMCSA, 2020; 2002). Commercial motor vehicle (CMV) drivers face all the routine risks of potential distractions as passenger drivers; however, they may need to engage in work-related communications while driving (e.g., routes, delivery scheduling, etc.). This may require communication between the CMV driver and their employer, shipper, receiver, etc. while in their vehicle. One method for this communication is via the cell phone.

Findings from three CMV driver naturalistic studies found no significant difference in the odds of a safety-critical event (SCE) between a handheld cell phone conversation and no cell phone conversation while driving a CMV (i.e., neutral factor). These same studies found that hands-free cell phone conversations significantly decreased the odds of a SCE (i.e., protective factor) compared to no cell phone conversation (Hammond et al., 2021; Hickman et al., 2012; Olson et al., 2009).

However, these studies combined different event types (e.g., sideswipe, head-on, etc.) in their analyses. It is possible cell phone conversations while driving have differential risk based on event type. Several studies have found that drivers are more likely to look at the forward roadway while conversing on a cell phone, which means hazards in a driver's forward view, such as read-end striking event types, would be more likely to be detected due to fewer off-road glances (Fitch et al., 2013; Hammond et al., 2021; Klauer et al., 2006; Olson et al., 2009; Victor et al., 2015). Does this greater attention to the forward roadway sacrifice attention to hazards on the side of the vehicle? And, if so, would they be detected in an analysis that combined event types? For example, rear-end event types were prevalent in naturalistic driving studies (Hammond et al., 2021; Dingus et al., 2016; Klauer et al., 2006; Olson et al., 2009). Analyses that combined event types would be skewed toward more prevalent event types.

In a preliminary analysis of the Strategic Highway Research Program 2 (SHRP2) dataset, Victor et al. (2015) found that drivers had a significantly lower odds of a rear-end striking crash or near-crash when engaged in a cell phone conversation compared to no cell phone conversation. Balint et al. (2020) used the SHRP2 dataset to evaluate the odds of multiple secondary tasks while driving. Multiple secondary tasks performed while driving significantly increased the odds of a run-off-road crash and rear-end striking crash by 4.09 and 6.94 times, respectively, compared to no secondary tasks. However, this analysis failed to isolate which specific combinations of secondary tasks increased the odds of these events. The current study evaluated the potential odds of a SCE during a cell phone conversation compared to no cell phone conversation stratified by event type in a sample of CMV drivers.

## 2. Method

The study used existing annotated data from Olson et al. (2009) and Hammond et al. (2021). Here we provide a summary of the methods, with a focus on the data used in the analyses. Although Hammond et al. (2009) was published in 2009, the data are still relevant today as CMV drivers continue to engage in hands-free and handheld cell phone conversations while driving. In addition, Hammond et al. (2021) included Bluetooth hands-free conversations that were not available when data were collected in Olson et al. (2009).

See Olson et al. (2009) and Hammond et al. (2021) for a detailed overview of the methods.

### *2.1 Overview of Datasets*

Olson et al (2009) included continuous (i.e., from ignition on to ignition off) naturalistic driving data from 203 CMV drivers in 55 instrumented trucks (all Class 8 tractor-trailers, gross vehicle weight rating exceeding 33,000 lbs.). The dataset included 4,452 SCEs and 19,888 random baselines. Hammond et al. (2021) included continuous naturalistic driving data from 172 CMV drivers in 182 instrumented trucks. The dataset included 2,363 SCEs and 7,880 random baselines.

These datasets included annotations derived from the SCEs, including: event type, driver ID, severity (crash, near crash, or crash-relevant conflict), and specific secondary tasks performed. Baselines included the same annotations as the SCEs; however, there were no annotations for event type and severity. Operational definitions for these variables can be found in Olson et al. (2009) and Hammond et al. (2020).

### *2.2 Data Stratification*

The data were stratified into seven event types in reference to the instrumented truck: (1) road departure, (2) rear-ending a stopped vehicle, (3) rear-ending a slower or decelerating vehicle, (4) side-swipe, (5) forward impact with a vehicle moving in the opposite direction (avoiding forward vehicle/object that resulted in an opposite direction SCE), (6) forward impact with a vehicle moving in the same direction (avoiding forward vehicle/object that resulted in a same direction SCE), pedestrian or pedacyclist, parked vehicle, fixed object, or construction barrier or construction cone, and (7) turning or crossing paths at an intersection. The event types were determined based on coding in the datasets corresponding to the "Accident Types" described in Olson et al. (2009). Excluded from the analysis were struck-by incidents in the "rear-end" and "forward impact with vehicle moving in same direction" categories, as well as incidents in these categories for which it was ambiguous as to whether the subject vehicle was striking or struck-by.

### *2.3 Analysis Approach*

Odds ratio (OR) estimates were calculated for handheld and hands-free cell phone conversation by event type. Baselines from drivers not involved in the event type were removed. The formula for calculating the OR is the cross product shown in Equation 1 (Agresti, 1996). Where n11 is the number of SCEs where the CMV driver was conversing on a cell phone while driving, n12 was the number of random baselines where the CMV driver was conversing on a cell phone while driving, n21 was the number of SCEs where the CMV driver was not conversing on a cell phone while driving, and n22 was the number of random baselines where the CMV driver was not conversing on a cell while driving.

The formula for calculating the confidence interval or CI is shown in equation 2. Where e is a constant and the base of natural logarithms, OR is the odds ratio, z is the z-score value corresponding to the chosen alpha (1.96 for a 95% CI), and SE is the standard error of the natural logarithm of the OR (Agresti, 1996). Significant odds ratio estimates greater than one "1.0" indicate the cell phone conversation increased the likelihood of a SCE, whereas significant odds ratio estimates less than one "1.0" indicate the cell phone conversation decreased the likelihood of a SCE. Non-significant odds ratio estimates indicate the cell phone conversation did not increase or decrease the likelihood of a SCE (Agresti, 1996; Mosteller, 1968; Woolf, 1955)

$$(n11 \times n22) \div (n12 \times n21) \qquad (1)$$

$$OR \times e \mp z \times SE^{OR} \qquad (2)$$

## 3. Results

Table 1 shows the counts of SCEs and random baselines, OR estimates, and 95% confidence intervals (CIs) by event type and conversation type. The authors compared the risk of a cell phone conversation with any other secondary task. For example, the OR calculations for the conversation type "handheld" included all instances where the driver was coded as having a handheld cell phone conversation and compared to all instances where the driver was coded with any other secondary (i.e., no handheld conversation, but potentially any other secondary task). Thus, the driver could have performed another secondary task category in this analysis.

There were nine significant findings in Table 1 (shown with an "*" in Table 1). Handheld cell conversations compared to no handheld conversations had a neutral odds ratio estimate regarding involvement in a rear-end striking event where the lead vehicle was stopped or slower/decelerating and forward impact (same direction) event types. Handheld cell phone conversations significantly increased the odds of involvement in a road departure that resulted in a SCE compared to no handheld conversations. The remaining cell phone conversation types by event type comparisons significantly decreased the odds of a SCE compared to no cell phone conversation. No analysis was performed for handheld conversations in forward impact (opposite direction) SCEs due to a cell count of zero.

## 4. Discussion

To our knowledge, this was the first analysis of the event-specific odds of a SCE regarding a cell phone conversation while driving a CMV. The results were largely consistent with other CMV studies showing that cell phone conversation, regardless of hands-free or handheld, did not significantly increase the odds of a SCE compared to no cell phone conversation (Hammond et al., 2021; Hickman et al., 2012; Olson et al., 2009). These results were supported by Victor et al's (2015) preliminary analysis of rear-end striking crashes in the SHRP2 dataset. In addition, Muttart et al. (2021) found no difference in brake response times to a lead vehicle in passenger car drivers when comparing drivers with and without a cell phone conversation.

## 5. Conclusions

Overall, a hands-free cell phone conversation while driving a CMV decreased the odds of a SCE compared to no cell phone conversation, regardless of event type. The results in this study support the existing Federal Motor Carrier Safety Administration regulations for CMV drivers, which allow for a hands-free conversation while driving a CMV (Federal

Register, 2011). These results support the Task Capability Interface Model (Fuller & Santos, 2002), and later referenced by Kinnear et al. (2007), which illustrates the relationship between user capabilities and overall task demands associated with driving (i.e., an increase in task demands does not necessarily mean that driver capabilities are exceeded).

## 6. Acknowledgments

The study from which these data were collected was funded by the Federal Motor Carrier Safety Administration under Contract DTMC75-07-D-00006. The Task Order Manager was Theresa Hallquist and the Contracting Officer's Technical Representative was Dr. Martin Walker. The opinions expressed in the paper are those of the authors and do not necessarily represent official positions of any government agency.

## References

Agresti, A. (1996). An Introduction to Categorical Data Analysis." Hoboken, NJ: John Wiley & Sons, Inc., ISBN 978-0-471-22618-5.

Bálint, A., Flannagan, C.A.C., Leslie, A., Klauer, S., Guo, F., & Dozza, M. (2020). Multitasking Additional to Driving: Prevalence, Structure, and Associated Risk in SHRP2 Naturalistic Driving Data. Accident Analysis and Prevention, 137, 105455, doi: 10.1016/j.aap.2020.105455.

Dingus, T.A., Guo, F., Lee, S., Antin, J.F., 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 113(10),636-2,641, doi: 10.1073/pnas.1513271113.

Dingus, T.A., Owens, J.M., Guo, F., Fang, Y., Perez, M., McClafferty, J., Buchanan-King, M., & Fitch, G.M. (2019). The Prevalence of and Crash Risk Associated with Primarily Cognitive Secondary Tasks. Safety Science 119, 98-105, doi: 10.1016/j.ssci.2019.01.005.

Federal Motor Carrier Safety Administration. (2022). Large Truck and Bus Crash Facts 2020. Washington, DC: Federal Motor Carrier Safety Administration.

Federal Motor Carrier Safety Administration. (2002). Report to Congress on the Large Truck Crash Causation Study. Washington, DC: 2. Federal Motor Carrier Safety Administration.

Fitch, G. M., Soccolich, S.A., Guo. F., McClafferty, J., Fang, Y., Olson, R.L., Perez, M.A., Hanowski, R.J., Hankey, J.M., & Dingus, T.A. (20130. The Impact of Hand-Held and Hands-Free Cell Phone Use on Driving Performance and Safety-Critical Event Risk. (Report # DOT HS 811 757)." Washington, DC: National Highway Traffic Safety Administration.

Fuller, R., & Santos, J.A. (2002). Psychology and the highway engineer. In Ray Fuller and Jorge Santos (Ed), Human Factors for Highway Engineers. Emerald Group Publishing Limited, 1-10.

Hammond, R.L., Soccolich, S.A., Han, S., Guo, F., Glenn, T.L., & Hanowski, R.J. (2021). Analysis of Naturalistic Driving Data to Assess Distraction and Drowsiness in Drivers of Commercial Motor Vehicles. Washington, DC: Federal Motor Carrier Safety Administration.

Hickman, J.S., & Hanowski, R.J. (2012). An Assessment of Commercial Motor Vehicle Driver Distraction using Naturalistic Driving Data. Traffic Injury Prevention 13(6), 566-574, DOI: 10.1080/15389588.2012.683841.

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

Mosteller, F. (1968). Association and Estimation in Contingency Tables. Journal of the American Statistical Association, 63, 1-28, doi: 10.2307/2283825.

Muttart, J., Dinakar, S., & Edewaard, D. "Drivers' Responses to Lead Vehicles: Thresholds for Triggering an Emergency Response, Age Differences, Crash Risks, and Influence of Secondary Task Engagement. SAE Technical Paper 2021-01-0898, doi: 10.4271/2021-01-0898.

Olson, R. L., Hanowski, R.J., Hickman, J.S., & Bocanegra, J. (2009). Driver Distraction in Commercial Vehicle Operations." Report No. FMCSA-RRR-09-042. Washington, DC: Federal Motor Carrier Safety Administration.

Victor, T., Dozza, M., Engstrom, J., Flannagan, C., Lee, J.D., & Markkula, G. (2015). Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk. SHRP 2 Report S2-S08A-RW-1. Washington DC: Transportation Research Board.

Woolf, B. (1955). On Estimating the Relationship Between Blood Group and Disease. Annals of Human Genetics, 19, 251-253, doi: 10.1111/j.1469-1809.1955.tb01348.x.

*Table 1 ORs for handheld (HH) and hands-free (HF) cell phone conversations by event type.*

| Event Type | Type | SCEs w/ Cell Phone | SCEs w/ No Cell Phone | Baselines w/ Cell Phone | Baselines w/ No Cell Phone | OR | 95% CI |
|---|---|---|---|---|---|---|---|
| Rear-End Stopped | HH | 1 | 144 | 398 | 10,341 | 0.180 | 0.025-1.293 |
| | HF | 3 | 142 | 1,005 | 9,775 | 0.205* | 0.065-0.646 |
| Rear-end Slower/ Decelerating | HH | 35 | 1,111 | 777 | 21,661 | 0.877 | 0.622-1.124 |
| | HF | 46 | 1,124 | 1,121 | 21,317 | 0.778 | 0.576-1.052 |
| Road Departure | HH | 133 | 3,637 | 572 | 19,296 | 1.233* | 1.018-1.495 |
| | HF | 70 | 3,700 | 981 | 18,887 | 0.364* | 0.285-0.465 |
| Forward Impact (Same Direction) | HH | 7 | 131 | 851 | 10,656 | 0.669 | 0.312-1.436 |
| | HF | 4 | 134 | 1,167 | 10,340 | 0.265* | 0.098-0.716 |
| Forward Impact (Opposite Direction) | HH | 0 | 85 | 846 | 4,900 | N/A | N/A |
| | HF | 2 | 83 | 1,136 | 4,610 | 0.098* | 0.024-0.398 |
| Sideswipe | HH | 18 | 1,081 | 790 | 23,244 | 0.490* | 0.306-0.785 |
| | HF | 39 | 1,060 | 1,179 | 22,855 | 0.713* | 0.515-0.987 |
| Turning | HH | 6 | 408 | 869 | 17,659 | 0.299* | 0.133-0.671 |
| | HF | 14 | 400 | 1,323 | 17,205 | 0.455* | 0.266-0.778 |

# Assessment of the Prevalence, Nature, and Risk of Distracted Driving Among Large Truck Drivers in Canada: A Naturalistic Driving Study

Rebecca Hammond[1], Lesheng Hua[1*], Patrick Byrne[2], David Edwards[1], Feng Guo[1], & Sheila G. Klauer[1]

[1] Virginia Tech Transportation Institute, 3500 Transportation Research Plaza, Blacksburg, VA 24060, USA
[2] Ministry of Transport of Ontario, 77 Wellesley Street West, Toronto, Ontario, Canada
[*]lhua@vtti.vt.edu

**Abstract: Commercial motor vehicle (CMV) drivers have a significantly increased risk of being involved in a safety-critical event (SCE) while engaged in cell phone-related tasks that require a driver's visual attention and physical manipulation. Although research on the prevalence of distraction behaviour in CMV drivers has increased over the past two decades, few studies have extensively investigated driver behaviour changes with respect to evolving cell phone capabilities. This study investigated current CMV driver interactions with cell phones and their effects on driving risk. Participants comprised 39 drivers employed at trucking fleets located in the province of Ontario, Canada. The Virginia Tech Transportation Institute installed data acquisition systems in 35 large trucks, some shared by multiple drivers, to record key driver behaviour information surrounding both SCEs and baseline driving events. Approximately 3,000 baseline control segments were randomly selected, and 735 SCEs were identified through a set of sensor trigger values. Results show that CMV drivers were engaging in secondary tasks in 47.5% or nearly half of all baseline epochs; 19.8% of all baseline epochs involved cell phone use. Hands-free talking/listening on the phone had the highest frequency but did not result in significant odds ratios. Drivers who had visual and/or manual interactions with a mounted cell phone had 5.80 times greater odds of being involved in an SCE compared to those not interacting with a mounted cell phone.**

## 1. Introduction

Commercial motor vehicle (CMV) drivers have a significant increased risk of being involved in an SCE while engaged in cell phone-related behaviours that require a driver's visual attention and physical manipulation (Olson et al., 2009; Hickman et al., 2010; Hammond et al., 2016; Hammond et al., 2021). Police collision reports in the United States and Ontario have indicated that approximately 15% of large truck collisions resulting in injury or fatality involved a distracted large truck driver (National Center for Statistics and Analysis, 2017; Byrne et al., 2020).

Previous studies have shown that tasks with high visual and manual requirements were dangerous potential distractions, such as texting and dialling a phone, reading and looking at maps from phone or navigational devices, and reaching for an object (i.e., cell phone, headset, sunglasses, and other objects; Olson et al., 2009; Hickman et al., 2010; Blanco et al., 2016; Hammond et al., 2021). Text messaging on a cell phone poses the highest risk while driving, whereas talking/listening on a hand-held or hands-free cell phone did not elevate the likelihood of being involved in a safety-critical event (SCE; Olson et al., 2009).

In Canada and the United States, regulations have been issued to restrict CMV drivers from reaching for or holding a mobile phone (76 Fed. Reg. 75470, 2011; Canadian Council of Motor Transport Administrators, 2018). Despite most drivers being aware of the risks associated with cell phone use while driving, 45% of drivers still reported using their phones (Claveria et al., 2019). Additionally, evolving cell phone technology leads to changes in driver behaviours. Therefore, this study aims to maintain an understanding of real-world driver cell phone use behaviours and investigate how drivers are currently interacting with cell phones while driving.

## 2. Method

### 2.1 Participants

This study involved 39 drivers employed at trucking fleets with home terminals located in the province of Ontario, Canada. They all held an Ontario Class A driver's license, which permits the operation of any heavy and commercial vehicle. The average age of drivers was 49.8 years, and their average length of CMV driving experience was 19.68 years.

### 2.2 Data Acquisition System

The VTTI-designed MicroDAS (see Figure 1) was installed in 35 large trucks to collect continuous kinematic and video data any time the vehicle was on and in motion in real-world environments. The kinematic and video data were used to identify and record key driver behaviour information surrounding both SCEs and baseline driving events.

***Fig. 1.*** *Data acquisition system.*

*2.3 Data Reduction*

A total of 1,528,618 kilometres of data were collected from all vehicles. The data reduction process involved creating triggers and utilizing VTTI software to detect kinematic thresholds that may lead to SCEs. Trained coders determined the validity of triggers, categorized SCEs, and coded various driver behavior variables. Seven hundred thirty-five SCEs were identified through a set of trigger values, and approximately 3,000 baseline control segments were randomly selected and stratified based on vehicle kilometres travelled. Secondary tasks were important variables, encompassing any behaviours in which drivers have been engaged just prior to the start of SCEs. The following table presents cell phone use tasks that were aggregated into several variables for the analysis (see Table 1).

***Table 1*** *Cell phone secondary tasks aggregated for analysis.*

| **Original secondary tasks** | **Aggregated secondary tasks** |
|---|---|
| Cell phone, browsing<br>Cell phone, texting<br>Cell phone, dialling hand-held<br>Cell phone, locating/reaching/answering | Hand-held phone visual/manual |
| Cell phone, holding | Hand-held phone manual |
| Cell phone, holding and glancing, hand-held | Hand-held phone visual |
| Cell phone, talking/listening, hand-held | Hand-held talking/listening |
| Cell phone, browsing hands-free mounted<br>Cell phone, dialling hands-free mounted | Mounted phone visual/manual |
| Cell phone, look at hands-free mounted<br>Cell phone, video calling hands-free mounted | Mounted phone visual |
| Cell phone, talking/listening, hands-free | Hands-free talking/listening |

There are two options for hands-free cell phone use. Drivers would mount a cell phone to the dash or centre stack or they would prop it in a cup holder for interacting with the screen. For the second option, they used a headset or Bluetooth earpiece for hands-free talking/listening.

## 3. Results

*3.1 Estimate Frequency of Secondary Tasks*

This analysis revealed that CMV drivers are engaging in secondary tasks in 47.5% of all baseline epochs. Twenty-two percent of baseline epochs were identified as some form of cell phone-related behaviour. Hands-free cell phone talking/listening has the highest frequency and percentage of baseline observations (13.35%) of all secondary tasks (see Table 2). Interacting with a mounted cell phone has the third highest frequency.

***Table 2*** *Frequency and percentage of cell phone use in baseline observations.*

| Secondary task | Percentage | Frequency |
|---|---|---|
| Hand-held phone visual/manual | 2.12 | 64 |
| Hand-held phone manual | 0.2 | 6 |
| Hand-held phone visual | 0.2 | 6 |
| Hand-held talking/listening | 0.13 | 4 |
| Mounted phone visual/manual | 1.36 | 41 |
| Mounted phone visual | 2.45 | 74 |
| Hands-free talking/listening | 13.35 | 403 |
| Other cell phone | 0.03 | 1 |

*3.2 Odds Ratios of Secondary Tasks*

Using the frequency of secondary task engagement during SCEs and baseline epochs, odds ratios were calculated to estimate the risk of SCE occurrence using a logistic regression model. Odds ratios are a comparison of the odds of SCE occurrence given a driver's secondary task behaviour compared to alert, non-distracted driver behaviour. A mixed effects logistic regression model was fit to the observations that satisfied the above conditions for each secondary behaviour. Drivers who had visual and/or manual interactions with a mounted cell phone had 5.80 times greater odds of being involved in an SCE compared to those not interacting with a mounted cell phone. This secondary task had the highest estimated odds ratio of all the cell phone-related behaviours. Additionally, talking with hands free is not associated with an increase in SCE occurrence (see Table 3).

***Table 3*** *Odds ratios and 95% confidence interval for cell phone use.*

| Secondary Task | Odds Ratios | Lower Confidence Limit | Upper Confidence Limit |
|---|---|---|---|
| Hand-held phone visual/manual | 3.66* | 2.27 | 5.91 |
| Hand-held phone manual | 2.01 | 0.38 | 10.60 |
| Hand-held phone visual | 3.72 | 0.85 | 16.30 |
| Hand-held talking/listening | 2.06 | 0.19 | 22.14 |
| Mounted phone visual/manual | 5.80* | 3.31 | 10.18 |
| Mounted phone visual | 1.18 | 0.56 | 2.51 |
| Hands-free talking/listening | 0.98 | 0.72 | 1.34 |

2

## 4. Discussion

Cell phone use continues to be a serious problem and could potentially worsen. Despite law restrictions, drivers still engaged in risky cell phone-related tasks. Regarding the significantly high odds ratios of hand-held phone visual/manual tasks, along with the insignificantly low odds ratio of hands-free talking/listening, these findings were consistent with previous studies (Olson et al., 2009; Hickman et al., 2010; Hammond et al., 2016; Hammond et al., 2021). On the other hand, the current manner of cell phone use is different than in previous studies. Drivers frequently mount their cell phone to the dash for browsing, dialling, and looking, especially on a long-haul trip. However, in contrast to hands-free talking/listening, mounting a cell phone is the riskiest behaviour, which is a unique finding. Other interesting results regarding secondary task engagements and their significance were not reported here.

Training programs and cell phone use policies for fleet trucking companies may need to be designed to better emphasize the importance of using hands-free technologies and/or to educate people on the risks of visual/manual tasks with both hand-held and mounted cell phone use when in complex driving environments.

## 5. Conclusions

CMV drivers were engaging in secondary tasks in 47.5% or nearly half of all baseline epochs; 19.8% of all baseline epochs involved cell phone use. The frequency of hands-free talking/listening on the phone was the highest but did not result in significant odds ratios. The odds ratio of mounted phone visual/manual cell phone use was the highest odds ratio. Drivers who had visual and/or manual interactions with a mounted cell phone had 5.80 times greater odds of being involved in an SCE compared to those not interacting with a mounted cell phone. These findings suggest that additional public service announcements/educational campaigns are needed to highlight the risks of any interaction (e.g. including mounted cell phone) with an electronic device. Additionally, in-vehicle display developers need to continue to improve voice command technologies to increase use as well as reduce physical interactions with screens. All stakeholders need to take the necessary steps to reduce the prevalence of tasks that take the driver's eyes off the forward roadway. Tasks which require drivers to look away from the forward roadway has repeatedly been shown to increase risk of crash involvement.

## References


Blanco, M., Hickman, J. S., Olson, R. L., Bocanegra, J. L., Hanowski, R. J., Nakata, A., Greening, M., Madison, P., Holbrook, G. T., & Bowman, D. (2016). *Investigating critical incidents, driver restart period, sleep quantity, and crash countermeasures in commercial vehicle operations using naturalistic data collection* (FMCSA-RRR-13-017). U.S. Department of Transportation.

Byrne, P. A., Rubin, F., Plonka, S., Monster, K., Dessau, E., & Elzohairy, Y. (2020). *Ontario's Large Truck Collision Causation Study: Final report*. Ontario Ministry of Transportation. https://www.library.mto.gov.on.ca/SydneyPLUS/Sydney/Portal/default.aspx?component=AAAAIY&record=8b67dfff-ac5d-4f54-bddb-8c9d39f419e7

Canadian Council of Motor Transport Administrators. (2018). *Distracted driving white paper*. https://www.ccmta.ca/web/default/files/PDF/CCMTA_Distracting_Driving_White_Paper_-_Revised_December_2018.pdf

Claveria, J. B., Hernandez, S., Anderson, J. C., & Jessup, E. L. (2019). Understanding truck driver behavior with respect to cell phone use and vehicle operation. *Transportation Research Part F, 65*, 389–401.

Fed. Reg., Vol. 76, No. 232, December 2, 2011, Docket No. FMCSA-2010-0096, Page 75470. Drivers of CMVs: Restricting the Use of Cellular Phones

Hammond, R. L., Hanowski, R. J., Miller, A. M., Soccolich, S. A., & Farrell, L. J. (2016). *Distraction and drowsiness in motorcoach drivers* (FMCSA-RRR-15-017). Federal Motor Carrier and Safety Administration, USDOT.

Hammond, R. L., Soccolich, S. A., Han, S., Guo, F., Glenn, T. L., & Hanowski, R. J. (2021). *Analysis of naturalistic driving data to assess distraction and drowsiness in drivers of commercial motor vehicles* (Report No. FMCSA-RRR-20-003). Federal Motor Carrier and Safety Administration, USDOT.

Hickman, J., Hanowski, R., & Bocanegra, J. (2010). *Distraction in commercial trucks and buses: Assessing prevalence and risk in conjunction with crashes and near-crashes* (Report No. FMCSA-RRR-10-049). Federal Motor Carrier and Safety Administration, USDOT.

National Center for Statistics and Analysis. (2017, March). *Distracted driving 2015* (Traffic Safety Facts Research Note. Report No. DOT HS 812 381). National Highway Traffic Safety Administration.

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009, September). *Driver distraction in commercial vehicle operations* (Report No. FMCSA-RRR-09-042). Federal Motor Carrier and Safety Administration, USDOT.

# Self-report survey of student driver engagement with distractions – the role of penalty points and experience

Terry C. Lansdown*, Elliot J. Kovanda, Liam Spence

Heriot-Watt University, Edinburgh, EH14 4AS
* t.lansdown@hw.ac.uk

**Abstract: 530 student drivers completed an anonymous survey over seven years. Findings related to penalty points, experience, and accidents, and their relationship to driver distraction. The survey collected demographic information, along with frequency and rated severity of engagement with distracting behaviours, and personality scale items. Questions considered 'internal to vehicle' and 'external to vehicle' distractions, for both 'work-related' and 'leisure' driving. Significant models predicted that i) as drivers' penalty points increase, they are more willing to engage with (internal to the vehicle) distracting behaviours; and ii) increasing driver experience predicted higher penalty points. Findings provide evidence of ongoing and repeated engagement with distracting behaviours.**

## 1. Introduction

Young drivers have repeatedly been shown to be one of the highest risk driving groups (RAC, 2009). Further, the current generation of young drivers have been exposed to smartphones for most of their lives, and many see them as an integral part of their lives. A considerable literature clearly demonstrates that unmanaged smartphone use while driving is remorselessly negative to road safety.

Previously published data from the survey instrument reported here (Lansdown et al., 2021) suggested that the internal to vehicle behaviours rated as most distracting were 'writing texts', 'internet use' and 'reading texts' while driving. For external to vehicle distractions, they were 'environmental conditions', 'unexpected objects or events', and 'animals behaving unexpectedly'. The most frequently undertaken in-vehicle behaviours were '(interactions with) adults', 'daydreaming', and 'eating, drinking or smoking'; while the external to vehicle ones were 'people (behaving normally), 'busy roads' and 'official signage'. The internal to vehicle distractions were reported to be relatively more distracting than external to vehicle ones. Respondents were found to experience more distractions during leisure driving than during work-related driving. It was not previously possible to report on the data regarding individual differences, and this paper presents these along with the associated research hypotheses.

The investigation had four hypotheses:

1. Higher penalty points will predict increased self-reported distracting behaviours.
2. Higher penalty points will predict greater accident involvement.
3. Greater driving experience will result in lower penalty points.
4. Greater driver experience will reduce accident involvements.

## 2. Method

### 2.1. Design

An anonymous online self-report survey was undertaken to collect the data presented here, based on Lansdown (2012). Collected data items included demographics, frequency and rated severity of engagement with distracting behaviours, along with personality scale items (the IPIP, Donnellan et al., 2006).

### 2.2. Procedure

Psychology students participated for course credit. After consent, respondents experienced five survey sections, i) demographics; ii) ratings of distracting behaviours, as listed in Table 1. For each item, participants responded with how distracting (1: not distracting – 5: very distracting); iii) how frequently they have experienced the item (I haven't done this while driving, daily, weekly monthly or yearly); iv) accident and near-miss history; and v) the IPIP battery, plus a field for any other comments and feedback.

Engagement with distracting behaviours was defined using an index calculated from the self-reported frequency of undertaking the distracting behaviours reported in Table 1. Ordinal values were assigned to generate a summative score for each respondent's activity, corresponding to '1' for yearly, '2' for monthly, '3' for weekly, and '4' for daily, for each behaviour, e.g., reading text messages on a weekly basis would accrue an item score of '3'. These values were added for each of the behaviours described to generate each respondent's distraction index.

### 2.3. Respondents

530 students responded to the survey during a seven-year data collection period between 2012 and 2018. All respondents were UK licensed drivers. In summary, 20% of drivers were male, average age was 20.6 (SD = 4.1) and on average they had 2.1 years of driving experience (SD = 3.2). Average mileage was 5.7 thousand miles per year (SD = 5.4). 96% of drivers had no penalty points. Of those with penalty points, 16 had three or less, five had between four and six points. 82.3% (436) of respondents had had no incidents in the last five years of driving. For those who with recent (withing the last five years) accident history, 14.9% (79) reported one, 2.3% (12) two, and 0.6% three incidents. For the 'at fault' incidents, one event was reported by 8.9% (49) of drivers, and two by 0.9% (5 drivers). A more detailed breakdown of demographic features and frequency of engagement data is presented in Lansdown et al. (2021).

**Table 1** Distractions and (average) ratings (1: not distracting – 5: very distracting).

| In-Vehicle | | External to Vehicle | |
|---|---|---|---|
| Write text | 4.16 | Environmental conditions | 3.84 |
| Internet (use) | 3.99 | Events (unexpected behaviours) | 3.58 |
| Read text | 3.58 | Animals (unexpected behaviours) | 3.56 |
| Handheld (device use) | 3.45 | People (unexpected behaviours) | 3.50 |
| (Navigation) destination entry | 3.22 | Other events | 3.29 |
| Other behaviour | 3.02 | Roads (complex) | 3.19 |
| Daydreaming | 2.88 | Roads (busy) | 2.99 |
| (Interaction with) children | 2.59 | Roadworks | 2.96 |
| Media player | 2.48 | Advertisements (dynamic) | 2.74 |
| Pets | 2.47 | Advertisements (static official) | 2.12 |
| Handsfree (device use) | 2.44 | Advertisements (static unofficial) | 2.06 |
| Headphones | 2.19 | Signs (official) | 1.91 |
| Eating, drinking or smoking | 2.10 | Signs (unofficial) | 1.86 |
| (Interaction with) adults | 1.92 | Animals | 1.77 |
| | | People | 1.51 |

## 3. Results

Data demonstrated heteroscedasticity during assumption checking. Therefore, weighted-least-squares (WLS) regressions were employed.

H1 proposed that those with higher penalty points would engage more with self-reported distracting behaviours. Significant models emerged from linear regressions for both work-related (including commuting; $t$ (1,491) = 6.487, $p < 0.0001$) and leisure-related driving ($t$ (1,514) = 8.912, $p < 0.0001$). For work-related, Penalty Points explained 28.1% of the variance in engagement with distracting behaviours, for leisure-related driving it was 36.6%. Thus, a one Penalty Point increase for work-related driving was predicted to increase engagement with (internal to vehicle) distractions (Distraction Index = Penalty Points * 2.086 + 23.537), and for leisure driving the same was also found (Distraction Index = Penalty Points * 1.808 + 27.625).

Greater numbers of penalty points were predicted to result in increased accident involvement in H2. A significant linear regression model emerged, however the WLS regression was found to be no longer significant ($t$ (1,528) = 1.825, $p = 0.069$). Consequently, the null is accepted for H2.

H3 predicts that greater driving experience will result in lower penalty points. The WLS regression retained a significant model ($t$ (1,527) = 4.228, $p <$ 0.0001). Experience explained 18.1% of the variance in Penalty Points. Contrary to H3, increased years of driving experience were found to predict increased Penalty Points (Experience = Penalty Points * 0.069 + -0.009). Thus, the null is accepted for H3.

H4 proposes that greater driver experience will reduce accident involvements. The model that emerged from linear regression was found to be non-significant ($F$ (1,527) = 2.78, $p < 0.096$). Therefore, the null was accepted for H4.

## 4. Discussion

Higher penalty points were predicted to increase engagement with distracting behaviours. Significant models emerged suggesting rejection of the null for H1. There seems some intuitive sense that those who have demonstrated themselves to be willing to engage with antisocial driver behaviours, e.g., excessive speed, may also be more prone to distractions.

Hypothesis 2 posited that higher penalty points will predict more accidents. There is robust meta-analytic support that penalty point systems reduce accidents, fatalities and injuries (Castillo-Manzano & Castro-Nuño, 2012). Thus, it was assumed that a significant relationship would be demonstrated in the data presented here. It seems that the positive effects of penalty point systems are only maintained for around eighteen months. It may be that the predominantly young respondents with penalty points investigated here; may have moderated their behaviours such that any effects from initially riskier behaviours were mediated. However, our lack of significant results for H2, preclude anything other than an acceptance of the null.

For Hypothesis 3 experience was predicted to result in lower penalty points. Findings were in opposition to our hypothesis. It may have been that the respondents from this survey, while still relatively new drivers, we relatively more experienced than those that have effectively just passed their driving test. Consequently, the predicted increase in penalty points may potentially be explained by exposure.

The final hypothesis was that greater driver experience will reduce accident involvements. As indicated above there is some evidence to support this hypothesis, in general, in the literature. Even for the relatively inexperienced group of respondents in the survey reported here, it was considered that consistent behaviours would be found. However, no significant relationship was established. As for H3 it could have been that our sample was in an in-between phase of their driving experience, no longer novices but not yet experienced.

## 5. Conclusions

A positive predictive relationship was found for drivers with penalty points and their engagement with distracting behaviours, for both work-related and leisure driving. No relationship was found for penalty points and accident involvement. Similarly, no relationship between experience and accidents was found either. Surprisingly, a significant positive relationship was found between experience and penalty points, contrary to expectations.

## 6. References

Af Wåhlberg, A. E. (2011). The accident-exposure association: Self-reported versus recorded collisions. *Journal of Safety Research*, *42*(2), 143–146.

https://doi.org/10.1016/j.jsr.2011.01.007

Castillo-Manzano, J. I., & Castro-Nuño, M. (2012). Driving licenses based on points systems: Efficient road safety strategy or latest fashion in global transport policy? A worldwide meta-analysis. *Transport Policy*, *21*, 191–201. https://doi.org/10.1016/j.tranpol.2012.02.003

Daigneault, G., Joly, P., & Frigon, J. Y. (2002). Previous convictions or accidents and the risk of subsequent accidents of older drivers. *Accident Analysis and Prevention*, *34*(2), 257–261. https://doi.org/10.1016/S0001-4575(01)00014-8

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The Mini-IPIP Scales: Tiny-yet-effective measures of the Big Five Factors of Personality. *Psychological Assessment*, *18*(2), 192–203. https://doi.org/10.1037/1040-3590.18.2.192

HESA. (2020). *HE student enrolments by subject of study 2014/15 to 2018/19*. https://www.hesa.ac.uk/data-and-analysis/students/table-9

Lansdown, T. C. (2012). Individual differences and propensity to engage with in-vehicle distractions - A self-report survey. *Transportation Research Part F: Traffic Psychology and Behaviour*, *15*(1). https://doi.org/10.1016/j.trf.2011.09.001

Lansdown, T. C., Kovanda, E. J., & Spence, L. (2021). Student driver propensity to engage with distractions – a self-report survey. *Transportation Research Part F: Traffic Psychology and Behaviour*, *81*, 650–660. https://doi.org/10.1016/J.TRF.2021.07.008

McCartt, A. T., Shabanova, V. I., & Leaf, W. A. (2003). Driving experience, crashes and traffic citations of teenage beginning drivers. *Accident Analysis and Prevention*, *35*(3), 311–320. https://doi.org/10.1016/S0001-4575(02)00006-4

Parmentier, G., Chastang, J. F., Nabi, H., Chiron, M., Lafont, S., & Lagarde, E. (2005). Road mobility and the risk of road traffic accident as a driver: The impact of medical conditions and life events. *Accident Analysis and Prevention*, *37*(6), 1121–1134. https://doi.org/10.1016/j.aap.2005.06.012

RAC. (2009). *RAC report on motoring 2009*. Royal Automobile Club. papers3://publication/uuid/98871383-D618-4077-AE4D-9AE25DAC4D6B

# Driver Profiles of Visual Manual Phone Engagement and The Contexts in Which They Occur

Rashmi P. Payyanadan [1], Josh Domeyer [2*], Linda Angell [3], Tina Sayer[4]

[1,3] Touchstone Evaluations, Inc., 81 Kercheval Ave., Ste 200, Grosse Pointe, MI 48236, USA
[2,4] Toyota Collaborative Safety Research Center, 1555 Woodridge Ave., Ann Arbor, MI 4810, USA
[*]joshua.domeyer@toyota.com

**Abstract: Phone use while driving, particularly visual-manual phone engagement (VMPE) like texting, dialing, and browsing, can significantly increase crash likelihood due to prolonged diversion of attention from the driving task. However, significant knowledge gaps remain regarding the circumstances surrounding VMPE, whether different types of VMPE are initiated in different contexts, and whether different driver types have distinct VMPE patterns. Additionally, there is limited understanding of drivers' VMPE across entire trips.**

**This study analyzed naturalistic driving data from 44 drivers collected over a three-week period. The data encompassed 557 full trips, driving kinematics, VMPEs, driving context, driver demographics, and psychosocial factors. The study aimed to predict the onset of VMPE as a function of driving context and driver-related variables.**

**Results showed that the strongest predictor of VMPE onset was the interaction between driving context, and a set of driver-related variables (i.e., driver's attitudes, injunctive norms, and perceived behavioural control toward speeding and phone use). Specifically, drivers with positive attitudes, injunctive norms, and perceived behavioural control toward speeding and phone use were more likely to engage in VMPE in certain driving contexts, such as, while waiting in traffic, during morning drives with low traffic, and during trips that had limited idling. VMPE driver profiles revealed three types: Consistent Context-Independent VMPE, Selective Context-Specific VMPE, and Minimal VMPE.**

**Together, these findings highlight the interplay between context and the driver's psychosocial factors, which significantly influence their VMPE. They emphasize the importance of considering driver characteristics when designing interventions specifically aimed at reducing VMPE.**

## 1. Introduction

Visual manual (VM) phone activities such as texting, dialling, and browsing are considered complex secondary tasks (C. Klauer et al., 2006). Engaging in complex secondary tasks significantly heightens the risk of crashes and near-crashes (Dingus et al., 2006; S. G. Klauer et al., 2006; Olson et al., 2009; S.G. Klauer et al., 2014). This heightened risk is attributed to how VM phone tasks involve extended periods looking away from the road compared to other types of phone tasks, causing drivers to allocate less attention to the road (T.W. Victor et al., 2005).

In 2021, eight percent of fatal crashes were reported as distraction-affected crashes (National Highway Traffic Safety Administration, 2023). Among those fatal crashes, 377 (12%) were linked to phone use, where at least one of the involved drivers was talking on, listening to, or engaged in phone activity at the time of the crash, resulting in 410 deaths. These findings underscore the importance of measures that can effectively reduce VM phone engagement while driving. However, tackling this issue requires a comprehensive understanding of *what* factors influence phone use while driving, *which* types of drivers are more inclined to engage with their phones, and *when* are they more likely to engage with their phones.

This study presents findings from a three-week naturalistic driving study (NDS) involving 44 drivers and 557 trips, aimed at understanding visual-manual phone engagements (VMPEs). The Comprehensive Driver Profile (CDP) analytical framework (Payyanadan & Angell, 2022) was implemented on the NDS data to evaluate the prevalence of VMPE across full trips, identify predictors of VMPE onset, and determine driver types and contexts based on VMPE.

## 2. Method

The NDS data used for the analysis in this study was collected by the University of Michigan Transportation Research Institute (UMTRI) and Collaborative Safety Research Center (CSRC). The dataset represents three weeks of driving and 557 trips taken by 44 drivers in 2018-19. Full trips were video-coded by UMTRI and validated (Molnar et al., 2021).

### 2.1 Driver demographics and psychosocial factors

The participant sample comprised of 15 younger (18-25 yrs.), 14 middle-aged (35-55 yrs.), and 15 older (65+ yrs.) drivers. Drivers completed seven surveys (Table 1).

*Table 1: List of surveys*

| Survey | Total survey items | Example |
|---|---|---|
| Demographics | 8 | age range, gender |
| Driving history | 9 | miles driven/week |
| Driving behavior questionnaire | 3 | Lapses, errors |
| Risky behavior engagement | 4 | Seatbelt use |
| Psychological well-being | 10 | Anger, anxiety |

| | | |
|---|---|---|
| Big five personality traits | 5 | Agreeableness |
| Theory of Planned Behavior (TPB) | 11 | Attitudes, norms |

### 2.2 *Driving kinematics and context*

Each of the 44 driver's personal vehicles were instrumented with a data acquisition system (DAS) and Mobileye (ME) system, consisting of a driver-facing and outward facing camera. See Table 3 in Appendix A for list of driving kinematics recorded by the DAS, and the derived and time-lagged kinematics variables used for the analyses. See Table 4 in Appendix A for list of driving context variables recorded by ME, and the derived and time-lagged driving context variables used for the analyses.

### 2.3 *Visual-manual phone engagement (VMPE)*

Full trips were video-coded for six types of phone activities: reaching/answering, talking/listening, hand-held texting/ browsing, dialing hand-held, holding, and other. For this study, the analyses focused specifically on VMPE activities: hand-held texting/ browsing, and dialing hand-held.

### 2.4 *Comprehensive Driver Profile (CDP) analytical framework*

The CDP framework involves: (a) conducting principal component analysis (PCA) to generate complex features representing interactions within driving behaviors, driving contexts, driver demographics, and driver psychosocial factors; (b) Random Forest (RF) algorithm for feature selection, prediction, and effect size extraction; and (c) *k*-means clustering for identifying types of drivers and driving contexts.

The CDP framework was implemented as follows. The PCA outcomes, in conjunction with raw and derived data, produced a total of 241 variables representing driving kinematics, driving context, driver characteristics, and psychosocial factors. Subsequently, all 241 variables were incorporated into the RF modelling to predict VMPE onset. Significant predictors of VMPE onset were then used to perform *k*-means clustering to identify types of drivers and driving contexts based on VMPE.

## 3. Results

### 3.1 *VMPE across drivers and full trips*

A total of 1,053 VMPE onsets were initiated by 44 drivers over a three-week period. Table 2 shows VMPE distribution by age group. Younger drivers exhibited a 25 percent higher rate of VMPE onsets/hour of driving compared to middle-aged drivers. Older drivers demonstrated the lowest rate.

*Table 2: VMPE by age group*

| Age group (yrs.) | Total VMPE | Total trip time (mins) | Number of VMPE onsets/hr of driving |
|---|---|---|---|
| 18-25 | 437 | 4,714 | 5.6 |
| 35-55 | 558 | 8,017 | 4.2 |
| 65+ | 58 | 5,818 | 0.6 |

### 3.2 *Predictors of VMPE onset*

The RF prediction model achieved an AUC of 0.76 in predicting VMPE onset, reflecting good discriminatory power. See Table 5 in Appendix A for prediction model outcomes. The results (Table 5) indicate four categories of VMPE predictors:

- Situations related to traffic
- Time of travel (Tot) on certain roads
- Positive attitudes, injunctive norms, and perceived behavioural control toward risk
- Willingness and intent to engage in risky behavior

Among the traffic-related situations, *low % of trip spent idling* had the highest RF variable importance, increasing the likelihood of VMPE by 20 percent. This could be attributed to drivers allocating most of their attention to the driving task thus far into the trip, and subsequently seeking a diversion upon encountering a period of idling. The substantial RF prediction variable importance of idling suggests that VMPE onset is primarily influenced by driving contexts associated with traffic situations. However, the relatively modest effect size of idling (20%) on the increase in VMPE suggests that while idling alone may not be the sole determinant of VMPE, it may serve as a significant situational trigger for certain drivers.

The two predictors that relate to Tot on certain roads, *morning drives on secondary, residential roads with little to no traffic*$_{pc}$ and *morning trip start time* both involved morning drives. This suggests that VMPE is most likely to occur during early morning trips, especially on secondary and residential roads, and when traffic is sparse.

The interaction effects emerged as the strongest predictors of VMPE onset. Specifically, drivers with positive attitudes, high injunctive norms, and high perceived behavioural control toward speeding, phone use, and NDRTs exhibited the highest effect size (> 50%) on the increase in VMPE. This suggests that drivers with more positive disposition toward speeding, phone use, and NDRT engagement have a higher likelihood of conducting VMPE when in traffic, during early morning drives, and while idling, compared to other drivers.

### 3.3 *Driver types based on their VMPE and the contexts in which they occur*

Clustering identified six driver types and four driving contexts associated with VMPE (Fig 1.). Among driver types, clusters 5 and 2 exhibited similar VMPE distribution, but differed in demographic composition (cluster 5, 60% females; cluster 2, 62% males). These two clusters represented drivers characterized as *Consistent Context-independent VM Phone Engagers* because they had the highest rate of VMPE, were relatively younger, and conducted VMPE regardless of context.

Drivers in clusters 1 and 3 represented *Selective Context-dependent VM Phone Engagers*. These drivers primarily differed in demographic composition (cluster 1, 75% males; cluster 3, 55% females) and typically engaged in VMPE only in specific contexts. While cluster 1 drivers mainly conducted VMPE during idling and high traffic

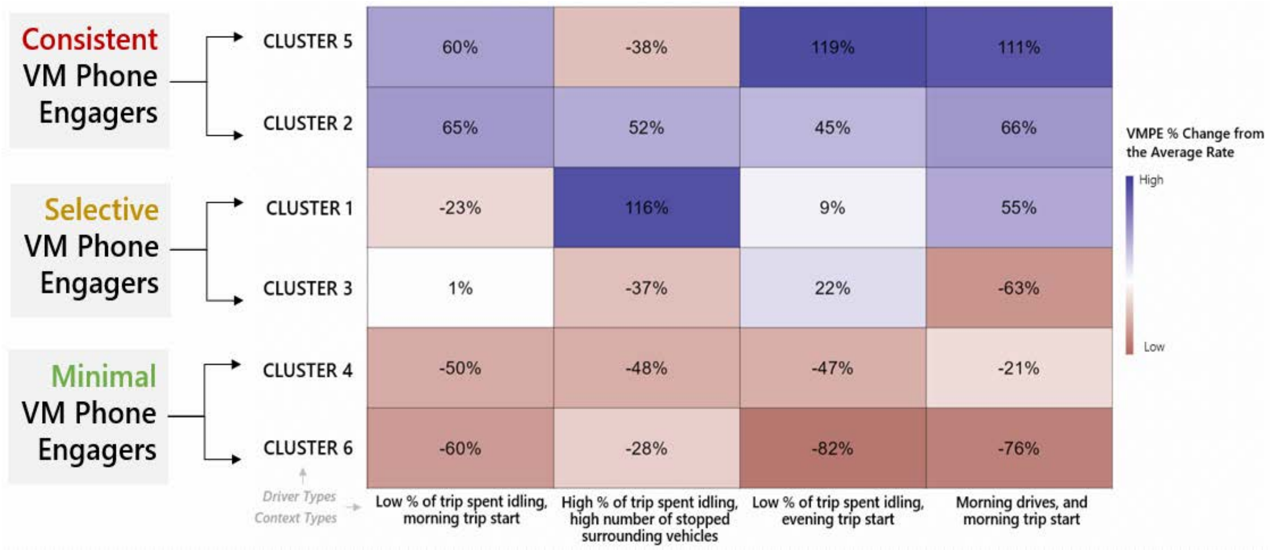



***Fig. 1.*** *Types of drivers and driving context. Percentages in each cell represents the VMPE percent change from the average VMPE rate for the driver and context cluster*

congestion; cluster 3 drivers mainly conducted VMPE during evening trips with low idling periods. Drivers in clusters 4 and 6 rarely conducted VMPE, and were classified as *Minimal VMPE Drivers*.

## 4. Discussion and conclusion

These findings highlight the interplay between driving context and the driver's psychosocial factors in influencing VMPE. They emphasize the importance of considering individual driver characteristics when designing interventions aimed at reducing VMPE.

## 5. Acknowledgments

This study was funded by the Toyota Collaborative Safety Research Center (CSRC). Special thanks to Sean Seaman from Touchstone Evaluations, Inc. for his invaluable feedback on this work.

## References

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., & Knipling, R. R. (2006). The 100-Car naturalistic driving study phase II – Results of the 100-Car field experiment. *National Highway Traffic Safety Administration*, *April*, No. HS-810 593. https://doi.org/DOT HS 810 593

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.*

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. In *NHTSA*.

Molnar, L. J., Eby, D. W., Zakrajsek, J. S., Kostyniuk, L. P., Zanier, N., LeBlanc, D. J., & Sayer, T. (2021). *Guidelines for Development of Evidence-Based Countermeasures for Risky Driving-Final Technical Report, Volume 2.*

National Highway Traffic Safety Administration. (2023). *Distracted driving in 2021 (Research Note. Report No. DOT HS 813 443)*.

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). *Driver distraction in commercial vehicle operations (No. FMCSA-RRT-09-042). United States. Department of Transportation. Federal Motor Carrier Safety Administration.*

Payyanadan, R. P., & Angell, L. S. (2022). A Framework for Building Comprehensive Driver Profiles. *Information*, *13*(2), 61.

S.G. Klauer, F. Guo, B.G. Simons-Morton, M.C. Ouimet, S.E. Lee, & T.A. Dingus. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. *New England Journal of Medicine*, *370*, 54–59.

T.W. Victor, J.L. Harbluk, & J.A. Engström. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*, 167–190.

## Appendix A

*Table 3: List of raw, derived, lagged driving kinematics variables*

| Variables | Unit of measure | Variable description |
|---|---|---|
| *Raw variables* | | |
| Trip start time | csec | Time when DAS turned on* |
| Trip end time | csec | Time when DAS turned off* |
| Trip distance$^{L}$ | miles | Distance travelled during a trip |
| Current speed$^{L}$ | m/sec | Vehicle speed from transmission |
| *Derived variables* | | |
| Total trips | Count | Total number of trips driven |
| Total trip time | sec | Total time from when the DAS turned on to when it turned off for a trip |
| Driving time thus far into trip$^{L}$ | Percent | Total duration of time lapsed in the trip since trip start time |
| Distance travelled thus far into trip$^{L}$ | Percent | Total distance covered in the trip since trip start time |
| Idling duration$^{L}$ | sec | Periods of time 5 secs after trip start time where the vehicle speed is < 1 mph |
| Idling frequency$^{L}$ | Count | Total number of *idling* events |
| Idling thus far into trip$^{L}$ | Percent | Total duration of time already spent *idling* at any given time in the trip |

*DAS on and DAS off reflects the time in centiseconds (csecs) when the ignition was turned on and off, respectively.
$^{L}$ reflects variables that were lagged by 2 secs, 4 sec, and 6 secs.

*Table 4: List of raw, derived, lagged driving context variables*

| Variables | Unit of measure | Variable description |
|---|---|---|
| *Raw variables* | | |
| Road type$^{L}$ | Values 0-13 | Road segments map-matched to Open Source Mapping (OSM): motorway, trunk, motorway link, primary, tertiary, secondary, residential, trunk link, tertiary link, primary link, secondary link, none, unknown |
| Posted speed limit$^{L}$ | m/sec | Way attribute in Open Source Mapping (not populated for every roadway) |
| Target ID | Values 0-9 | Unique ID assigned to each object tracked by the ME system in the driving environment |
| Target type$^{L}$ | Values 0-4 | Identification of the type of object being tracked by the ME system: car (0), truck (1), motorcycle (2), bicycle (3), pedestrian (4) |
| Target status | Values 0-6 | Identification of the status of the object being tracked by the ME system: undefined (0), standing (1), stopped (2), moving (3), oncoming (4), parked (5), unused (6) |
| Closest-in-path Vehicle$^{L}$ (CIPV) | Count | Time when the current headway of the preceding vehicles in the driver's lane and adjacent lane is too short to be safe where the host speed must be between 2.2 and 45 m/s (~5 and ~100 mph) |
| Range rate | m/sec | Relative longitudinal velocity of each object tracked by the ME system in the driving environment |
| *Derived variables* | | |
| Total trips | Count | Total number of trips driven |
| Traffic density$^{L}$ | Count | Total number of targets (moving, standing, and stopped cars and trucks) in the driving environment at any given moment in time. |
| Average surrounding vehicle speed$^{L}$ (svs) | m/ sec | Velocity or rate of movement of the moving vehicles (0, 1, 2, ..., *n*) in the immediate vicinity of the subject vehicle (sv). |
| Distance travelled thus far into trip$^{L}$ | Percent | Total distance covered in the trip since trip start time |
| Idling duration$^{L}$ | sec | Periods of time 5 secs after trip start time where the vehicle speed is < 1 mph |
| Idling frequency$^{L}$ | Count | Total number of *idling* events |
| Idling thus far into trip$^{L}$ | Percent | Total duration of time already spent *idling* at any given time in the trip |

$^{L}$ reflects variables that were lagged by 2 secs, 4 sec, and 6 secs.

*Table 5: Predictors of VMPE onset*

| Features | Importance scores | Main Effects: Effect size | Main Effects: % Effect (relative to mean overall VMPE onset) | Main Effects: Directionality |
|---|---|---|---|---|
| % of trip spent idling $_{\textit{6 seconds ago}}$ | 1 | 0.373 | 20.1% | - |
| Attitude, injunctive norms, and perceived behavioural control toward NDRT engagement $_{pc}$ (TPB NDRT) | 0.77 | 0.074 | 12.4% | + |
| Morning drives on secondary, residential roads with little to no traffic $_{pc}$ | 0.71 | 0.116 | 13.5% | + |
| Speeding is safe, pleasant, high perceived susceptibility to be pulled over speeding, others do not think it is okay to speed $_{pc}$ (TPB Speeding) | 0.70 | 1.106 | 56% | + |
| Attitude, injunctive norms, and perceived behavioural control toward phone engagement $_{pc}$ (TPB Phone) | 0.54 | 0.727 | 35.3% | + |
| Sensation seeking, self-esteem $_{pc}$ (TPB Psychosocial) | 0.52 | 0.127 | 20.5% | - |
| Morning trip start time | 0.40 | 0.303 | 17.2% | + |
| Intend to speed, low past speeding behavior $_{pc}$ (TPB Speeding) | 0.39 | 0.010 | 3.6% | . |
| Surrounding vehicle speed (stopped) $_{\textit{12 seconds ago}}$ | 0.36 | 0.877 | 42.1% | + |
| Sensation seeking, self-efficacy $_{pc}$ (TPB Psychosocial) | 0.32 | 0.356 | 15.8% | - |

| Interacting features | | Interaction Effects: Effect size | Interaction Effects: % Interaction effect (relative to mean VMPE onset) | Interaction Effects: Directionality |
|---|---|---|---|---|
| Speeding is safe, pleasant, high perceived susceptibility to be pulled over speeding, others do not think it is okay to speed $_{pc}$ (TPB Speeding) | Surrounding vehicle speed (stopped) $_{\textit{12 seconds ago}}$ | 0.874 | 66.8% | ++ |
| | Attitude, injunctive norms, and perceived behavioural control toward phone engagement $_{pc}$ (TPB Phone) | 0.724 | 60.5% | ++ |
| | % of trip spent idling $_{\textit{6 seconds ago}}$ | 0.373 | 56.7% | +- |
| | Sensation seeking, self-efficacy $_{pc}$ (TPB Psychosocial) | 0.342 | 54.7% | +- |
| | Attitude, injunctive norms, and perceived behavioural control toward NDRT engagement $_{pc}$ (TPB NDRT) | 0.075 | 52.8% | ++ |
| | Morning trip start time | 1.106 | 52.8% | ++ |
| | Sensation seeking, self-esteem $_{pc}$ (TPB Psychosocial) | 0.127 | 52.3% | +- |
| | Morning drives on secondary, residential roads with little to no traffic $_{pc}$ | 0.115 | 52.0% | ++ |
| Surrounding vehicle speed (stopped) $_{\textit{12 seconds ago}}$ | Attitude, injunctive norms, and perceived behavioural control toward phone engagement $_{pc}$ (TPB Phone) | 0.876 | 49.8% | ++ |
| | % of trip spent idling $_{\textit{6 seconds ago}}$ | 0.375 | 43.5% | +- |

*pc* represents the principal components outputted by the PCA in the CDP framework. Principal component variables represent linear combinations of the driving kinematics, driving contexts, driver characteristics, and psychosocial features.

# Temporal proximity and events duration affects change detection during driving

**Vasiliki Kondyli[1,*], Mehul Bhatt[2]**
[1]Jagiellonian University, Poland
[2]Örebro University, Sweden
Email: *vasiliki.kondyli@uj.edu.pl

**Failure in change detection in the surrounding environment while driving is attributed among other things to the number of incidents and the density of them occurring along the task as this is directly related to an increase in cognitive load. Here, we investigate the role of time proximity between events on the detection performance during a naturalistic driving task in a virtual simulation. Participants performed a change detection task while driving, in which we systematically manipulated the time difference between changes and we analysed the effect on detection performance by the driver. Our research demonstrates that events occurring simultaneously deteriorate detection performance (in terms of detection rate, and detection time), while performance improves as the temporal gap increases. Moreover, the outcomes suggest that the duration of an event affects the detection of the following one, with better performance recorded for very short or very long duration events and worse for medium duration events between (5-10 sec). These outcomes are crucial for driving assistance and training, considering the detection of safety-critical events or efficient attentional disengagement on time from irrelevant targets.**

## 1 Temporal proximity & event perception

Driving is a continuous task that requires a combination of cognitive processes among which is high-level visual processing in order to navigate, keep situation awareness, and track moving objects and relevant changes. Failure in detecting changes in the surrounding environment can be attributed to environmental complexity, the relevance of the event to the task, the characteristics of targets, as well as the temporal proximity between events. Temporal proximity between events is connected to temporal load and can lead to compromised perception along a continuous everyday task.

A growing body of research suggests that fine-grained event perception can be insensitive to brief temporal disturbances, meaning that events occurring with temporal delays of milliseconds up to a few seconds might be treated by many parts of the visual-cognitive systems as equivalent and so rapid succession of events leads to an almost universal degradation of detection performance. Specifically, according to research on dual-task interference (Pashler 1994; Raymond, Shapiro, and Arnell 1992), when two targets are presented in a time window of less than 100 msec, humans fail to encode the stimuli as two separate events Shallice 1964; VanRullen and Koch 2003, and similarly at temporal proximity of 100-500 msec, observers failed to report which stimulus was the first or second to appear, an effect known as the attentional blink (Sheppard et al. 2002; Raymond, Shapiro, and Arnell 1992). Additionally, event segmentation theory (Zacks et al. 2007), suggests that temporal sequence between short events in a several-seconds window may be represented by default and can be immediately perceived (James 1982). As confirmed by Pöppel 2009 and Fairhall, Albi, and Melcher 2014, conscious activities are integrated within 2-3 seconds windows, however, the task is getting more difficult in longer time windows. The effect of time proximity on event perception along the time window of a few seconds has not been thoroughly tested, leaving open questions on event perception, working memory capabilities, and the role of attentional blink as a cognitive strategy rather than a resource limitation (Wyble, Bowman, and Nieuwenstein 2009).

In the present study, we explore the hypothesis that temporal proximity between events affects detection performance and leads to adjustments in gaze behaviour along the course of events. Using the change detection paradigm (Simons and Levin 1998; Martens 2011; Kondyli et al. 2023), this research examines the effects of time proximity in a change detection task designed and embedded within a naturalistic everyday driving experience (implemented in virtual reality). Specifically, we ask: Do people miss more

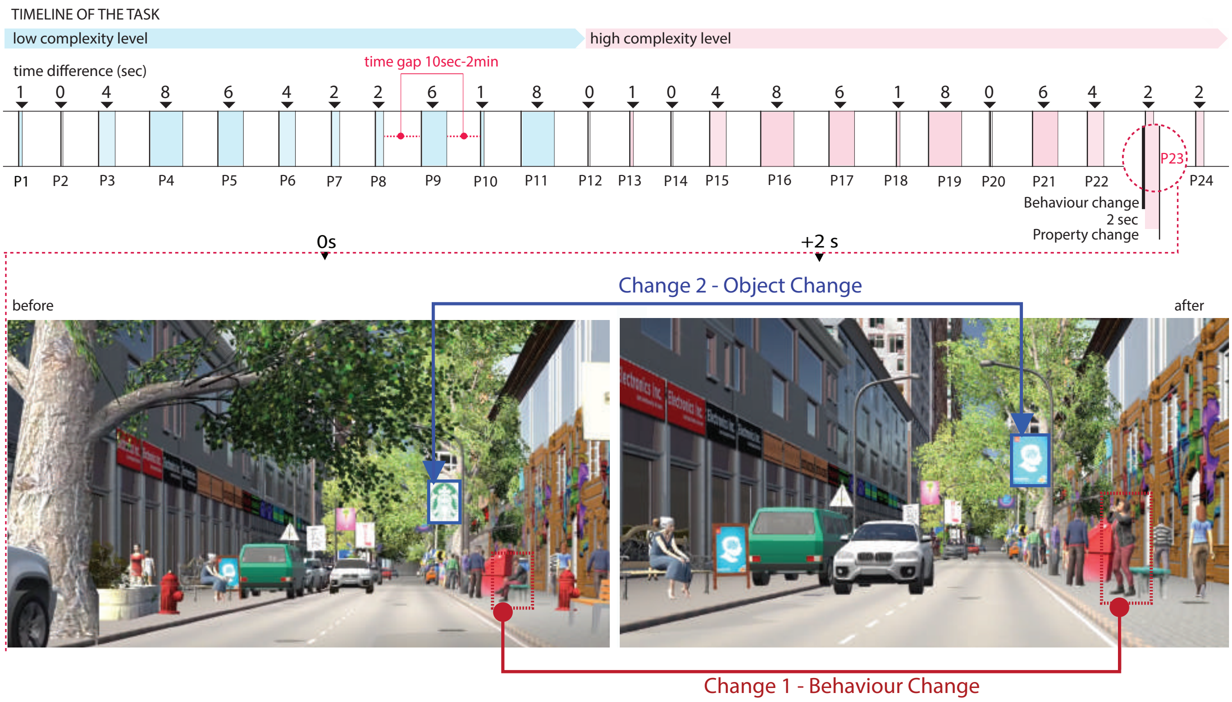



Figure 1: Structure of the experimental session: The horizontal line represents the timeline of the experiment and each vertical line represents one change occurring at a designated time point along the timeline. Two pairs of changes are illustrated in detail before and after the change occurs, including a behaviour change first and an object change following.

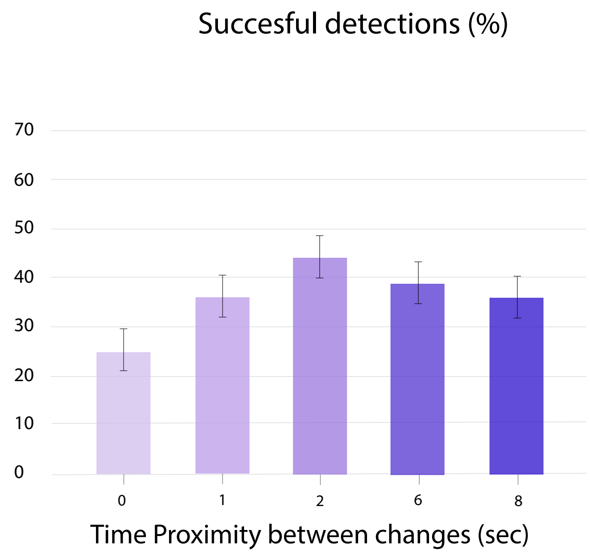
changes when they occur close in time between each other? What is the performance cost for changes occurring closely after a long event? Is there a time proximity threshold for effective detection of successive changes? We expect that shorter time gaps will result in worse detection performance for the second change and that longer duration of one change will delay attentional disengagement.

## 2 Method

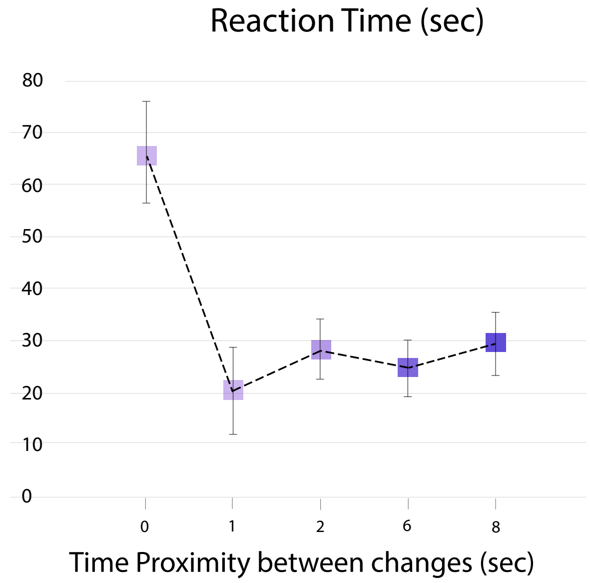
In a naturalistic behavioural study in VR, participants drove in an urban environment towards a designated destination, while they were instructed to detect changes in the behaviour of agents or changes in objects of their surrounding environment. The 72 changes were equally distributed in 36 pairs, in which a behaviour change is encountered first and an object change follows. We systematically manipulate the temporal proximity between the changes, ranging between 0 to 8 seconds (time gap of 0, 1, 2, 6, or 8 sec), and the duration of a change between 1 and 25 seconds (Fig. 1).

The selection of time gaps was based on previous studies on attentional blink, perceived duration, and event boundaries, suggesting that people need approximately 180-240 msec to detect visual stimuli and perceive duration (Jain et al. 2015; Efron 1970), more than 500 msec to distinguish between stimulus (Sheppard et al. 2002; Raymond, Shapiro, and Arnell 1992), 2-3 sec to integrate an activity or event (Zacks and Tversky 2001; Swallow, Kemp, and Candan Simsek 2018). As the literature provides different perspectives on the time gaps that affect perception, we combine the different perspectives and define accordingly the test levels between a minimum, at 0 sec, and a maximum, at 8 sec. Therefore, this range of time gaps makes it possible to test the previous theories on the amount of time needed to register and detect high-level events.

Considering the nature of the changes, we take the diversity of typical driving events into consideration and provide a number of interactive scenarios as the change detection task involving different agents (e.g., pedestrians, cyclists, kids, teenagers, older adults, people in wheelchairs) and street objects (e.g., parked cars, bus stops, sings). We recruited 80 participants (between 17-45 years old) who completed the task in 45-50 minutes. We collected multimodal behavioural data including gaze, head movements, and driving behaviour as well as detection performance analysis (detection rate and detection time) for the specific changes detection task.




(a) Detection rate for all groups of time proximity.




(b) Reaction time for all groups of time proximity.

Figure 2: Detection performance of object changes (2nd change), with respect to the time proximity from the 1st change. All error bars are standard errors.

## 3 Results

We examined the detection rate and the reaction time (RT) for the successfully detected object changes in relation to the time proximity from the behaviour changes. Time proximity serves as an independent variable (predictor) with five levels, where the time gap between the changes is one of 0, 1, 2, 6, 8 sec. A one-way between-participants ANOVA reveals a significant effect of time proximity on the change detection for all levels, $F(5, 474) = 14.286$, $p$<.001, with participants performing significantly worse when two changes happened simultaneously.

One-way ANOVA of RTs for the subset of changes that were detected (on average 36.6% of all object changes), showed a significant effect of time proximity on RT, $F(5, 926) = 4.440$, $p = .001, d =$. Participants were significantly slower at detecting an object change that happened at the same time as a behaviour change (0 sec level, M = 1.63, SE = 0.91) than in the case where there was a time gap of 1 to 8 sec between the two changes (Fig. 2b).

Overall, the results of detection performance, as well as the RT, suggest that detection performance for the second change was significantly compromised when this change was performed simultaneously (time proximity of 0 sec) to a behaviour change, with a probability of 24.8% to be detected, and an average RT of 1.633 sec. For the rest of the groups of time proximity, a recovering trend was observed, with the best performance recorded in the condition of 2 sec time proximity (with a probability of 43.8% to detect the change in an average of 1.306 sec).

Moreover, we analysed the duration of all behaviour changes among participants and we summarised the average time for each event. The 36 behaviour changes were experienced as events of 1 sec to 25 sec. We group the changes with respect to duration in three groups (short, medium, long). The detection analysis shows that medium-duration events interfere more with the detection of the following change than short and long events (Fig. 3). This result indicates that short and long events are related to more efficient coordination of attention, meaning that people disengage their attention on time, which allows for new targets be be detected.

## 4 Discussion

The study reveals that time proximity between events matters for efficiently perceiving them, and that the duration of an event affects the perception of the following one. Surprisingly, the duration of an event also matters as engaging attention for a longer time to an event may lead to disruptions in the detection of the following event. Examining these results from the viewpoint of *perceptual load theory* (Lavie and De Fockert 2003), we suggest that close time

| Duration of 1st change | up to 5 | 5 - 10 | 10 + |
|---|---|---|---|
| % of 2nd changes detected | 47.3 | 39.2 | 56.7 |

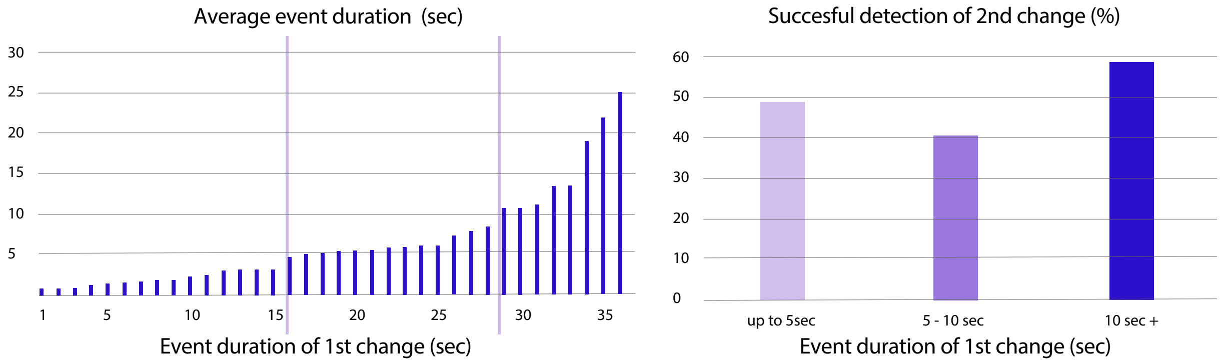



Figure 3: Analysis of the average duration of a behaviour change and its effect on detecting the following object change.

proximity between events may increase cognitive load, leading to disruptions in the efficient detection of changes. The low detection performance for changes that happen at the same time indicates the perceptual and sensory limitations of humans, however, the performance radically improves with an additional second between events. Further work will be necessary to identify the threshold of time gaps required for better detection performance and if performance can reach a plateau. Additionally, the analysis of event duration in relation to performance shows the importance of attentional disengagement from a target on time in order for people to have available attentional resources required for the upcoming targets along a continuous task.

Knowledge about the limitations and capabilities of human high-level visual processing in driving is crucial for driver's assistance systems as well as for training. For example, training for professional drivers should include monitoring and assessing changes in the behaviour of other vulnerable road users (e.g., kids, older adults), anticipating crossing behaviours in busy urban areas, responding to rapid changes of events, and keeping situation awareness in highly dynamic urban environments, monitoring blind spots, etc.

## References

Pashler, H (1994). "Dual-task interference in simple tasks: data and theory." eng. In: *Psychol Bull* 116.2, pp. 220–244. ISSN: 0033-2909 (Print); 0033-2909 (Linking).

Raymond, J. E., Shapiro, K. L., and Arnell, K. M. (1992). "Temporary suppression of visual processing in an RSVP task: An attentional blink?" In: *Journal of Experimental Psychology: Human Perception and Performance* 18.3, pp. 849–860.

Shallice, T. (1964). "The detection of change and the perceptual moment hypothesis". In: *British Journal of Statistical Psychology* 17, pp. 113–135.

VanRullen, R and Koch, C. (2003). "Is perception discrete or continuous?" In: *Trends in Cognitive Science.* 7.5, pp. 207–213.

Sheppard, D. M., Duncan, J., Shapiro, K. L., and Hillstrom, A. P. (2002). "Objects and Events in the Attentional Blink". In: *Psychological Science* 13.5, pp. 410–415.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., and Reynolds, J. R. (2007). "Event perception: a mind-brain perspective." eng. In: *Psychol Bull* 133.2, pp. 273–293. ISSN: 0033-2909 (Print); 1939-1455 (Electronic); 0033-2909 (Linking).

James, W. (1982). *Psychology: A Briefer Course.* New York: Henry Holt and Company.

Pöppel, E. (2009). "Pre-semantically defined temporal windows for cognitive processing." In: *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* 364, pp. 1887–1896.

Fairhall, S. L., Albi, A., and Melcher, D. (2014). "Temporal integration windows for naturalistic visual sequences." In: *PloS one* 9.7.

Wyble, B., Bowman, H., and Nieuwenstein, M. (2009). "The attentional blink provides episodic distinctiveness: sparing at a cost." eng. In: *J Exp Psychol Hum Percept Perform*

35.3, pp. 787–807. ISSN: 0096-1523 (Print); 0096-1523 (Linking).

Simons, D. J. and Levin, D. T. (1998). "Failure to detect changes to people during a real-world interaction". In: *Psychonomic Bulletin & Review* 5.4, pp. 644–649.

Martens, M. H. (2011). "Change detection in traffic: Where do we look and what do we perceive?" In: *Transportation Research Part F: Traffic Psychology and Behaviour* 14.3, pp. 240–250. ISSN: 1369-8478.

Kondyli, V., Bhatt, M., Levin, D., and Suchan, J. (2023). "How do drivers mitigate the effects of naturalistic visual complexity? On attentional strategies and their implications under a change blindness protocol". In: *Cognitive Research: Principles and Implications* 8.1, p. 54.

Jain, A., Bansal, R., Kumar, A., and Singh, K. (2015). "A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students." eng. In: *Int J Appl Basic Med Res* 5.2, pp. 124–127. ISSN: 2229-516X (Print); 2248-9606 (Electronic); 2229-516X (Linking).

Efron, R. (1970). "The minimum duration of a perception". In: *Neuropsychologia* 8.1, pp. 57–63. ISSN: 0028-3932.

Zacks, J. M. and Tversky, B. (2001). "Event structure in perception and conception". In: *Psychological Bulletin* 127.1, pp. 3–21.

Swallow, K. M., Kemp, J. T., and Candan Simsek, A. (2018). "The role of perspective in event segmentation". In: *Cognition* 177, pp. 249–262. ISSN: 0010-0277.

Lavie, N. and De Fockert, J. W. (2003). "Contrasting effects of sensory limits and capacity limits in visual selective attention". In: *Perception & Psychophysics* 65.2, pp. 202–212.

# Driver Attention Insights using Hybrid Naturalistic-Controlled Research

Joel M. Cooper

MEA Forensics Engineers & Scientists, 23281 Vista Grande Drive, Laguna Hills CA, USA
Joel.cooper@meaforensic.com

**This research examines driver secondary task engagement while using Level 2 vehicle automation through a hybrid controlled-naturalistic study spanning 6-8 weeks. Consistent with instructions, drivers used automation over 70% of the time. System warnings increased over time, suggesting complacency. The study also examined fatigue, arousal, and secondary tasks. Results highlight the importance of a strong control conditions in naturalistic research.**

## 1. Introduction

Automated vehicles (AVs) have the potential to significantly reduce traffic fatalities by addressing human error (Iden & Shappell, 2006). However, the passive monitoring role required by Level 2 (L2) systems, such as Adaptive Cruise Control and Lane Keep Assist, has raised concerns regarding increased driver disengagement and distraction, potentially due to the monotony of automated driving (SAE, 2018). This abstract synthesizes key findings from a study by Cooper et al. (2023), which investigated the impact of L2 vehicle automation on driver distraction and inattention.

Cooper et al. (2023) employed a hybrid controlled-naturalistic methodology to observe driver interactions with L2 automation over a 6-8 week period. The study examined several factors, including the influence of automation on driver arousal and fatigue (Matthews et al., 2019), and the likelihood of drivers engaging in non-driving related tasks, which could impair their ability to effectively monitor the system (Dunn et al., 2021).

This abstract focuses on the findings specifically related to driver distraction and inattention. By highlighting these aspects, the aim is to contribute to the broader understanding of the safety implications associated with the deployment of AV technologies. The insights provided seek to inform future research and development efforts in this field.

## 2. Method

Participants (N=30, 12 females, 18 males, aged 18-55) met eligibility criteria of a valid US driver's license, no recent at-fault accidents, no prior Level 2 automation experience, and a daily 20+ minute interstate commute.

Video data from 6-8 weeks of naturalistic driving were analyzed and coded for key behaviors. Participants were instructed to use L2 Automation as often as they felt comfortable for their daily commute. This created an "Automation: YES" condition with two resulting naturalistic conditions: drivers chose to use automation (Automation-L2) or they chose not to use automation. See Figure 1 below for additional information. Note that the naturalistic control comparison came from days where participants could use automation but chose not to and the experimental control came from the one day each week when participants were told not to use automation.

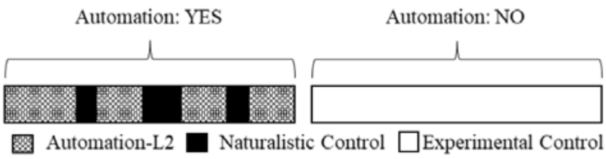



Figure 1. Schematic representation of the naturalistic and experimental control driving days representing when participants were allowed to use automation (Automation: Yes) and when they were not allowed to use automation (Automation: No – Experimental control).

(Naturalistic Control). Additionally, participants were instructed not to use automation one day each week, with video from these segments forming the "Experimental Control" condition.

Data reported in this abstract form a small subset of a larger effort which also included survey data (see Sanbonmatsu et al., 2023), and controlled on-road evaluations (see McDonnell et al., 2023).

Vehicles: Five commercially available models with Level 2 automation were used: Tesla Model 3 (n=6 participants), Tesla Model S (n=8), Cadillac CT6 (n=1), Volvo XC90 (n=9), Nissan Rogue (n=6). Participants were randomly assigned to one vehicle.

Cameras: Rosco Dual-Vision XC4 cameras were positioned in each vehicle just under the rear-view mirror, these returned video of the forward road and the driver. An auxiliary camera recorded the vehicle state display (location varied by model). Data was stored on Rosco and Transcend SD cards.

Video Coding: BORIS software (Friard & Gamba, 2016) was used for frame-by-frame video analysis, enabling pre-specification of behaviors of interest. Results were output to .csv files detailing behavior, location, and start/stop times.

Procedure: Video footage was restricted to participants' interstate commutes and combined into daily files. A two-pass blinding process ensured coders were unaware of experimental conditions (e.g., Automation L2, Naturalistic Control, and Experimental Control) during initial behavior coding. Coders received extensive training, and inter-rater reliability checks were conducted on at least 40% of videos. A comprehensive dictionary defined key behaviors: automation use, system warnings, driving demand, driver arousal (fatigue and fidgeting), and secondary task engagement (including modality and interface details).

BORIS video coding software was used to code driver behaviors and generate a .csv file of observations for analysis. We used R to transform this data into a time-series format, with behaviors as columns and binary indicators for task state. This structure enabled flexible recombining and collapsing of behaviors for various analyses. Linear mixed-effects models (using R's lmerTest library) were employed to account for repeated measures and missing data. Participant ID and AM/PM drive were included as random intercepts. Session and Condition were predictor variables, with likelihood ratio tests and pairwise comparisons used to determine significance.

## 3. Results

Data Overview. Video data from 30 participants yielded 670 total videos (353 Naturalistic, 317 Baseline). After excluding instances of automation use on Baseline days (due to task misunderstanding), 291 Baseline videos remained. Participant availability varied across study weeks. Of the available videos, redundancy coding was high (308 double-coded, 76 triple-coded, 4 coded by 4 individuals), with results averaged. Analysis focused on one coded Naturalistic day and one Experimental Control day per week, yielding 297 hours total. Over half of this (161 hours) captured Naturalistic driving, with participants using Level 2 automation for 25-99% of that time (124 hours).

Automation Usage. This research explored two questions about Level 2 automation usage over time: whether experience affected activation frequency and whether it influenced re-engagement time after disengagement. A mixed-effects model was used, with Week as a fixed effect, and Subject and AM/PM drives as random effects. Results indicated that Week did not significantly predict usage frequency or reengagement time. This suggests that participants' patterns of automation use, in terms of activation and re-engagement, remained consistent throughout the 6–8-week observation period (See Figure 2).

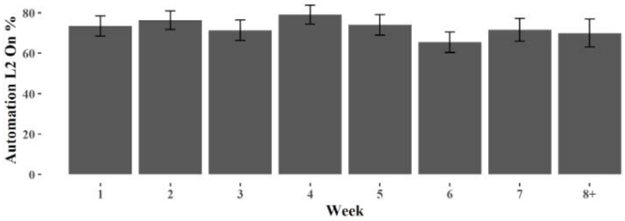



Figure 2. Automation L2 Use over time.

System Warnings and Driving Demand (See figure 3 below). This research addressed two questions about Level 2 automation misuse: whether system warning frequency changed over time and whether automation use varied based on driving conditions. System warnings (due to insufficient steering input or lack of forward gaze) ranged from 0.03-1.93 per minute. A mixed-effects model (Week as fixed effect, Subject and AM/PM drives as random) revealed that warning frequency increased over time. This suggests drivers may become less attentive with increased experience. To analyze automation use in relation to driving demand, poor conditions (traffic, weather, construction, etc.) were coded. Traffic impairment was most common. Another mixed-effects model showed automation use decreased as driving demand increased from Low to Moderate to High. This indicates drivers were aware of road demands and adjusted automation use accordingly.

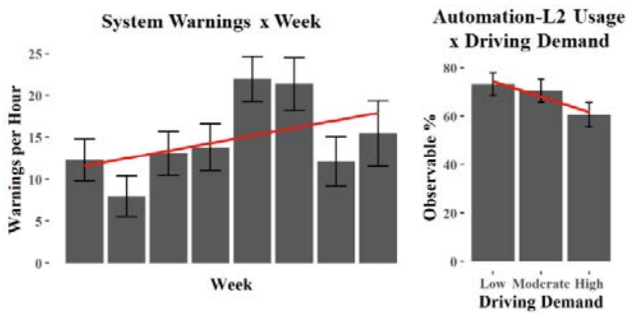



Figure 3. System warnings over time (Left fugre) and Automation use by driving demand (Right figure).

Secondary Task Engagement. This research explored how Level 2 automation affects the frequency of secondary task engagement (e.g., radio, texting). Results showed a main effect of Condition on overall task engagement, radio listening, and texting. Drivers engaged in more secondary tasks when using automation (Automation-L2) compared to manual driving (Naturalistic Control), but not compared to when automation use was prohibited (Experimental Control). Texting also increased over the study period, primarily in the Automation-L2 condition.

## 4. Discussion

The primary aim of this research was to better understand how the use of L2 automation affects driver secondary task engagement. We used a hybrid approach that married elements of naturalistic driving studies with a true experimental control condition. This structure facilitated a nuanced exploration of driver behavior in both controlled and more spontaneous driving conditions, focusing on patterns of automation use, response to system warnings, driver arousal (measured through fatigue and fidgeting signs), and engagement in secondary tasks. Findings reveal that drivers used L2 automation over 70% of the time during their commutes, a consistency echoing similar research findings and suggesting a comfortable reliance on these systems. Interestingly, the frequency of system warnings increased with time, indicating a possible relaxation in drivers' monitoring over the automation period, with usage patterns seemingly unaffected by varying driving demands. This observation aligns with previous studies indicating that drivers adjust their engagement with automation based on the perceived safety and roadway conditions. Such insights into automation usage and drivers' responsiveness to system warnings contribute significantly to understanding the operational dynamics of L2 automated vehicles in real-world settings.

While an increase in secondary task engagement was observed, the nature and potential safety impact of these behaviors varied, highlighting the importance of context in assessing the implications of automation on driver distraction. These findings challenge and refine our understanding of how L2 automation influences driver behavior, offering valuable insights for the development of safer automated driving systems.

## 5. Conclusions

This abstract illustrates results from a unique hybrid naturalistic research design with true experimental control condition developed by Cooper et al., (2023). Results illuminate the intricate balance between driver distraction, inattention, and reliance on Level 2 vehicle automation.

## 6. Acknowledgements

This research was funded by a generous grant from the AAA Foundation for Traffic Safety.

The presentation of this work at DDI is made possible by support from MEA Forensics Engineers and Scientists.

This research was funded by a generous grant from the AAA Foundation for Traffic Safety. The presentation of this work at DDI is made possible by support from MEA Forensics Engineers and Scientists.

## 7. References

Cooper, J. M., Crabtree, K. W., McDonnell, A. S., May, D., Strayer, S. C., Tsogtbaatar, T., ... & Strayer, D. L. (2023). Driver behavior while using Level 2 vehicle automation: a hybrid naturalistic study. *Cognitive research: principles and implications,* 8(1), 71.

Dunn, N. J., Dingus, T. A., Soccolich, S., & Horrey, W. J. (2021). Investigating the impact of driving automation systems on distracted driving behaviors. *Accident Analysis & Prevention,* 156, 106152.

Iden, R., & Shappell, S. A. (2006, October). A human error analysis of US fatal highway crashes 1990–2004. *In Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 50, No. 17, pp. 2000-2003). Sage CA: Los Angeles, CA: Sage Publications.

Matthews, G., Neubauer, C., Saxby, D. J., Wohleber, R. W., & Lin, J. (2019). Dangerous intersections? A review of studies of fatigue and distraction in the automated vehicle. *Accident Analysis & Prevention,* 126, 85-94.

SAE International (2018). Surface Vehicle Recommended Practice: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J3016). Warrendale, PA: Society of Automotive Engineers.

# Effect of Conflicting Information on Driving Behaviour in a Lane Change Task

Nikita Rajendra Sharma [1*], Jai Prakash Kushvah [2], Gerhard Rinkenauer [1]

[1] Ergonomics Department, Leibniz Research Centre for Working Environment and Human Factors at Dortmund University, Ardeystr. 67 44139 Dortmund, Germany
[2] Institute of Psychology, TU Dortmund University, Dortmund, Emil-Figge-Str. 50 44227 Dortmund, Germany
*Sharma@ifado.de

**Abstract: Preparatory information during driving has a significant influence on the driver's behaviour, as it supports the smooth execution of driving manoeuvres. However, conflicting, irrelevant information from a different sensory modality can significantly interfere with the initiation and execution of driving manoeuvres. To investigate how preparatory information and conflicting information affect the initiation and execution of driving manoeuvres, we used a lane change task in a simulated driving environment. Subjects were either informed in advance about the direction of a lane change or received no information. The target cue about when and in which direction the lane change should begin was presented simultaneously visually and acoustically. In one condition, the participants had to respond only to the visual target cue and in the other condition only to the acoustic target cue. The multimodal target information could be congruent or incongruent. The diffusion model for conflict tasks (DMC) was used as a theoretical framework for conflicts in information processing. The results showed that preparatory information helped participants to react faster and perform the lane change more effectively. During the multimodal presentation of the target cue, the visual cues were more effective. In addition, conflicting information, as predicted by the DMC, had a negative effect on reaction time and steering performance, especially when participant had to follow the auditory target cue. The response conflicts could not be compensated by preparatory information. The results are discussed in the context of the DMC.**

## 1. Introduction

Driving requires a continuous integration of sensory information from the environment and the vehicle in order to control current actions and prepare for upcoming events. Digital driver assistance systems such as navigation devices, traffic sign recognition or lane change assistants also provide advanced information about the current and future driving situation via visual, acoustic and haptic channels. The issue of response preparation has been frequently addressed in basic research (Brunia & van Boxtel, 2000; Müller-Gethmann et al., 2000; Requin et al., 1991). Rosenbaum (1980) provided the movement preparation technique, which aims to assess covert preparatory processes within the motor system. This technique has been used in many different basic experiments (Müller-Gethmann et al., 2000) as well as for the analysis of more complex rotational movements (Anson et al., 2000). Hofmann and colleagues used the preparation paradigm to investigate the effect of advanced information in lane change tasks (Hofmann, Rinkenauer & Gude, 2010 and Hofmann & Rinkenauer, 2011, 2013). Different information via different modalities ensures robust perception (Ernst & Bulthoff, 2004) and information processing as long as the different sources of information are congruent (Ngo & Spence, 2010, Green & Gierke, 1984; Cao et al, 2010; Sun, Y. 2016 and Lundkvist & Nykanem, 2016). However, if the information is contradictory (incongruent), this can lead to distraction or conflicts in information processing. This applies not only to the processing of relevant information aimed at achieving a goal, but also to the involuntary processing of irrelevant information (Morre, 2010).

In the current study, we investigated the extent to which relevant and irrelevant information from different modalities (audio and visual) about an upcoming lane change manoeuvre influences driving performance and how preparatory information can compensate for the conflict between contradictory action information causing distraction. The theoretical framework used was the Diffusion Model of Conflict (DMC) by Ulrich et al., 2015, which assumes that a single accumulation process combines information from controlled and automatic processes and determines a response based on the combined information (see also Logan, 1980, Logan and Zbrodoff, 1979).

We hypothesised that anticipatory response preparation would improve reaction times (RT) and steering wheel dynamics. In the context of DMC, it is expected that conflicting information will lead to longer RT and poorer steering performance. Furthermore, we expect that prepared responses are less susceptible to conflicting information than unprepared ones. Finally, we expect irrelevant visual information will interfere more with acoustic instructions than vice versa.

## 2. Method

### 2.1 Participants

Twenty young participants with valid driving license (age range 20–35 years; $M_{Age} = 24.40 \pm 3.27$ years, $M_{Driving\ experience} = 6.10 \pm 3.14$ years) took part in the experiment.

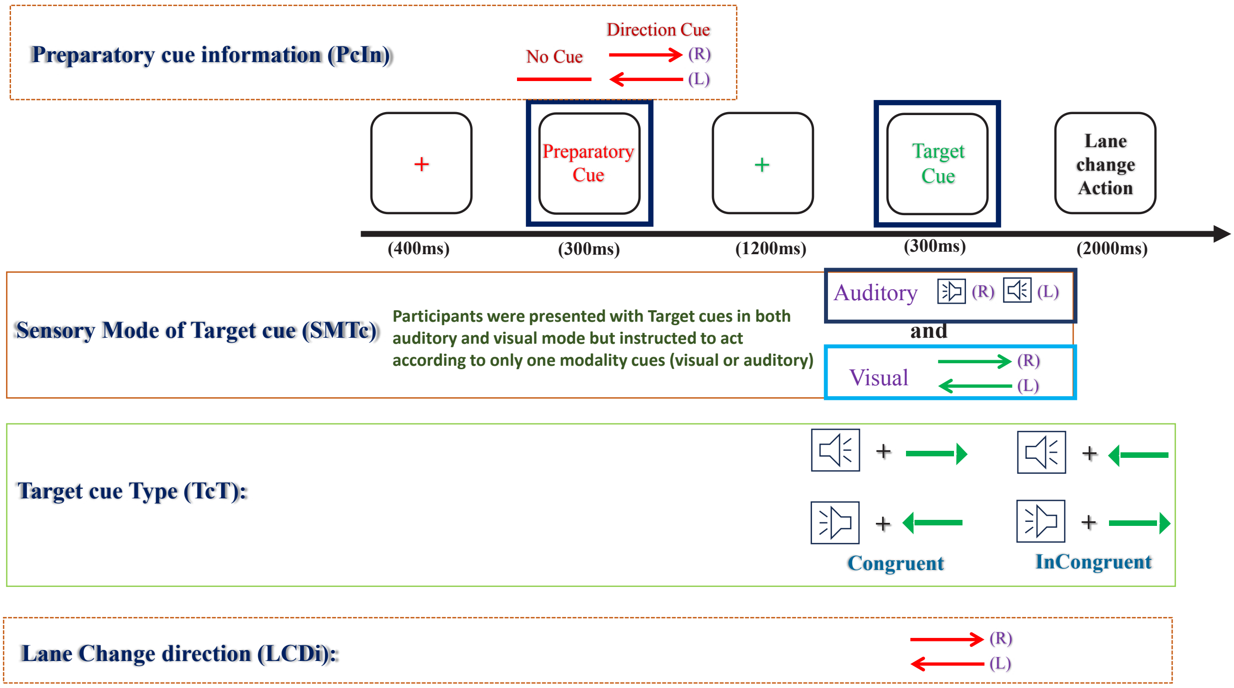



***Fig. 1.*** *Schematic representation of Lane Change Task showing the task trail run and experimental conditions.*

### *2.2 Driving task*

The current experimental paradigm corresponds to the previously used pre-cuing paradigm for LCT (Hofmann, Rinkenauer and Gude, 2010; Hofmann and Rinkenauer, 2013; Sharma, Kushvah and Rinkenauer, 2022). The task was extended to include a multimodal (visual and auditory) target cue (response signal). The target cues were always presented simultaneously, but subjects were asked to respond only to the instructed visual or auditory cue and to ignore the information of the other modality, analogous to the Stroop or Simon task (Lu and Proctor, 1995 and Boudreau, 2021). Figure 1 shows a schematic representation of the paradigm.

### *2.1 Procedure*

Participants received 10 minutes of practical training on the task. They were instructed to keep their vehicle in the middle of the road and to respond to the target cue as accurately and as quickly as possible. The experiment was conducted in two consecutive sessions with a total of 360 trials (10% of catch trials, cf. Buckolz and Rodgers, 1980). In one session, participants were instructed to respond to visual target cues only and in the other session, to acoustic target cues only. The order of the two sessions was counterbalanced across participants. Both sessions consisted of both congruent (when both target cues indicated the same direction of lane change) and incongruent (opposite) trials.

### *2.2 Statistical Analysis*

Mixed-factor repeated-measures ANOVAs were performed to assess the effects of the within-subjects factors including *Sensory Modality of Target cue* (SMTc: Visual vs. Auditory), *Target cue Type* (TcT: Congruent vs. Incongruent), *Preparatory cue Information* (PcIn: NoCue vs. Direction) and *Lane Change Direction* (LCDi: Right vs. Left) and a between-subjects factor of *Gender* (Male vs. Female) on RT and steering wheel angles (A1 and A2).

## 3. Results

Results showed that advance information reduced RT and improved steering wheel angle control of A1 and A2 [Figure 2 (a), (b) and (c)]. Faster RTs and reduced steering wheel angle A1 were observed in the trials where participants were asked to respond according to the visual target cues [Figure 3 (a) and (b)]. In addition, the congruent multimodal target cues showed shorter RTs and more effective motor control (A2) than incongruent target cues [Figure 4 (a) and (b)]. Thus, congruent multimodal target cues facilitated the lane change manoeuvres, whereas incongruent target cues created conflict and disrupted automated driving behaviour.

Furthermore, the results provided novel insights into the effects of pre-cue and multimodal target cues through several interaction effects. First, RT remained unaffected for both congruent and incongruent trials when participants were asked to follow the visual target cues, whereas RT increased significantly in the incongruent condition when participants were asked to follow the auditory target cues [Fig.5 (a), *SMTc* X *TcT*: RT (visual, congruent) ≈ RT (visual, incongruent) ≈ RT (auditory, congruent) < RT (auditory, incongruent)]. Thus, conflict was mainly produced when the incongruent target cues were present and the auditory target cues were required to be followed. Such an effect was further reflected, in the immediate steering wheel action (A1), i.e. A1 was higher during the incongruent target cue presentation when participants followed the auditory cues [Fig.5 (b), *SMTc* X *TcT*: A1(auditory, congruent) < A1(auditory, incongruent)]. Second, providing advance preparatory driving cues helped participants during the response preparation overall (reduced RTs) and congruence of target cues contributed significantly [Fig.6 (a), *PcIn* X *TcT*: RT (no-cue, incongruent) > RT (no-cue, congruent) > RT (direction, incongruent) > RT (direction, congruent)]. Third, preparatory direction cues were only useful for male participants in improvising their steering wheel control (A2) [Fig.6 (b), *PcIn* X *Gender*: A2(no-cue, male) > A2(direction, male)].

***Table 1*** *ANOVA results assessing the relationship between experimental main effects and LCT measures*

| LCT Measure | Experimental effects | F | P | $\omega^2$ |
|---|---|---|---|---|
| RT | PCIn | 41.55 | <.001 | 0.35 |
| | TcT | 52.34 | <.001 | 0.15 |
| | SMTc | 12.65 | .002 | 0.13 |
| | SMTc X TcT | 41.51 | <.001 | 0.11 |
| | PCIn X TcT | 6.78 | 0.02 | 0.01 |
| A1 | PCIn | 31.75 | 0.001 | 0.64 |
| | SMTc | 4.32 | 0.05 | 0.19 |
| | SMTc X TcT | 5.64 | 0.03 | 0.24 |
| A2 | PCIn | 8.38 | 0.01 | 0.01 |
| | TcT | 4.42 | 0.05 | 0.003 |
| | PCIn X Gender | 4.47 | 0.05 | 0.005 |

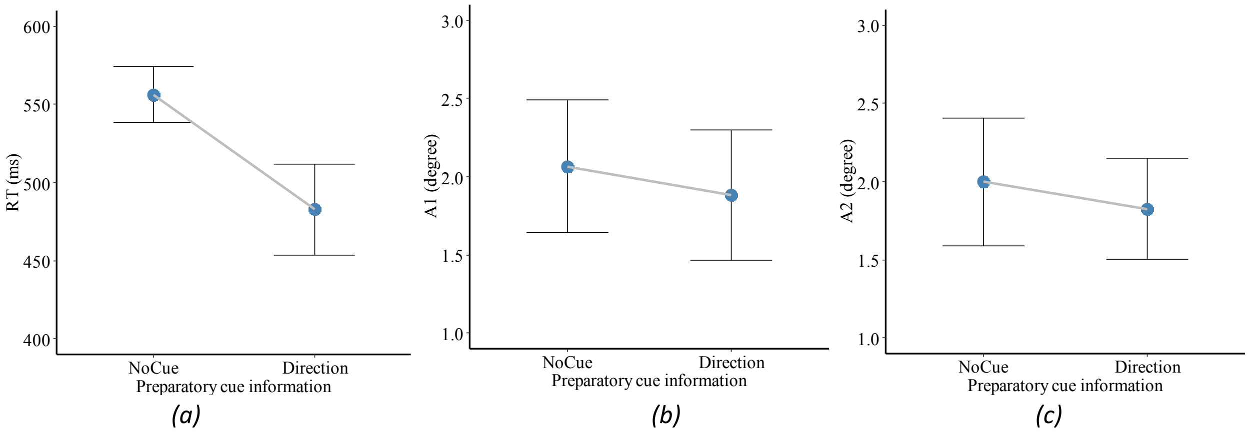



***Fig. 2.*** *Plots showing the mean comparision (with ±95% confidence interval (CI)) for (a) RT (b) A1, and (c) A2 as a function of prepratory cue information (PCIn).*

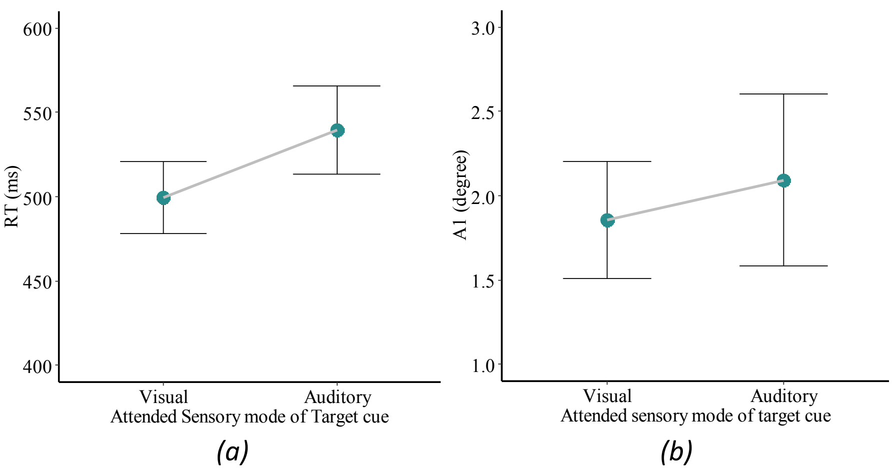
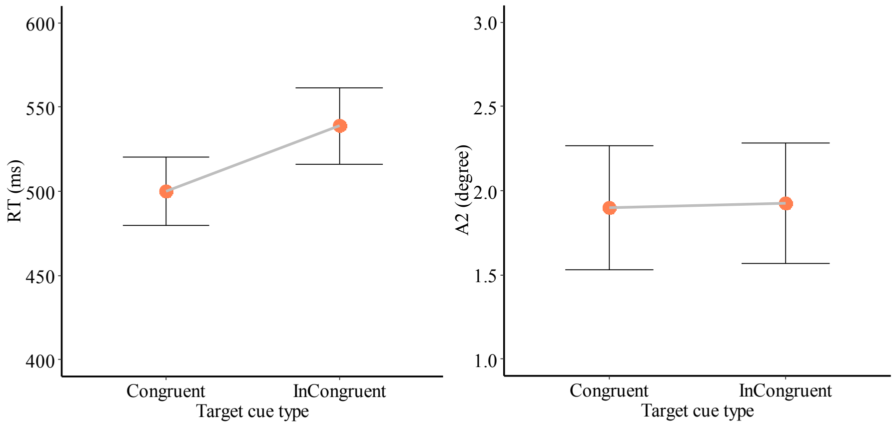
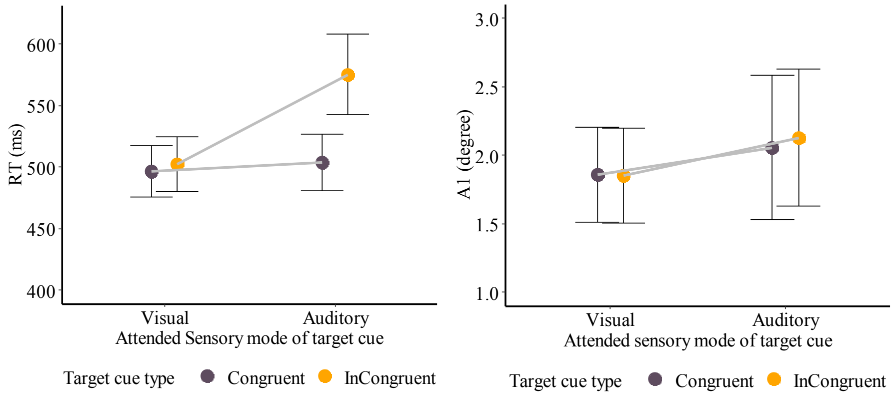



***Fig. 3.*** *Plots showing the mean comparision (with ±95% confidence interval (CI)) for (a) RT and (b) A1 as a function of sensory mode of target cue (SMTc).*



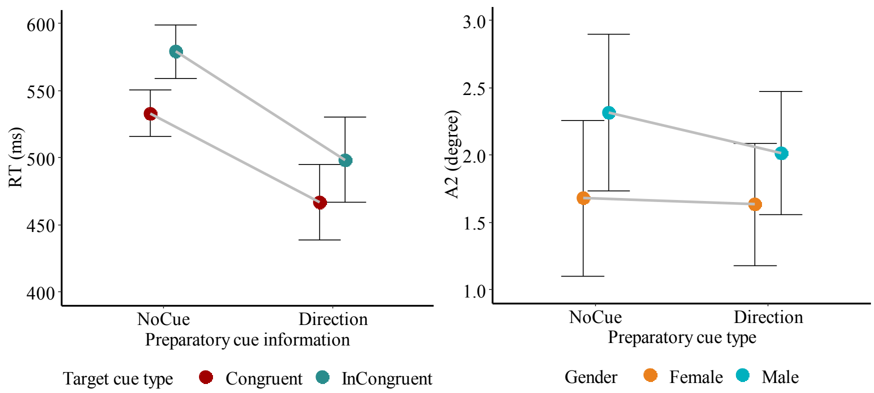

*(a)* *(b)*

***Fig. 4.*** *Plots showing the mean comparision (with ±95% confidence interval (CI)) for (a) RT and (b) A2 as a function of target cue type (TcT).*




*(a)* *(b)*

***Fig. 5.*** *Plots showing the mean comparision (with ±95% confidence interval (CI)) for (a) RT and (b) A1 as a function of attended sensory mode of target cue (SMTc) and target cue type(TcT).*




*(a)* *(b)*

***Fig. 6****. Plots showing the mean comparision (with ±95% confidence interval (CI)) for (a) RT as a function of prepratory cue information (PCIn) and target cue type (TcT), and (b) A2 as a function of prepratory cue information (PCIn) and Gender.*

## 4. Discussions and Conclusions

In our study, we consider distraction as a cognitive conflict and designed our empirical investigation accordingly. Our preliminary results suggest that irrelevant congruent (incongruent) information improves (worsens) driving performance as expected. Preparatory information before a lane change helped participants to change lanes faster and more smoothly than without preparatory information. The preparation effect found is consistent with existing research findings (Sharma, Kushvah & Rinkenauer, 2022; Hofmann, Rinkenauer & Gude, 2010 and Hofmann & Rinkenauer, 2013). Visual target cues were generally responded to more quickly than auditory target cues (Lundkvist & Nykanem, 2016). The results also show that the effect of response priming in congruent trials was more pronounced for visual target cues than for auditory target cues. This could be due to the fact that only visual pre-cues were used. The results are generally consistent with the predictions of the DMC that subjects have only limited success in suppressing automatic processes of incongruent irrelevant information. However, our hypothesis that preparatory information could compensate for the conflict between contradictory action information could not be confirmed. This would have been expected on the basis of the DMC, as prepared information processing processes start closer to the reaction criterion than unprepared processes (Ulrich et al. ,2015). Informative for the DMC is the finding that the conflict was greatest, and had the strongest negative effect on the lane change manoeuvre, when the reaction to the auditory target stimulus was required and the visual target stimulus was oriented in the opposite direction, which was also reflected in the immediate steering wheel action (A1). Analysing the steering reaction dynamics beyond the RTanalysis provides information on how conflicts in information processing affect the post RT action processes and helps to gain a deeper understanding of the mechanisms of conflictual information processing while driving.

## 5. Acknowledgments

The authors would like to thank Mr. Hanno Mussmann for his technical support and Ms. Larissa Achtermann for her support in the data collection.

## References

Anson, J. G., Kötter, R., & Wickens, J. R. (2000). Parameter precuing and motor preparation. *Motor control*, *4*(2), 221-231.

Boudreau, A. F. (2021). *Stimulus-Response Compatibility and the Simon Effect in a Simulated Driving Cockpit*. California State University, Long Beach.

Brunia, C. H. M. & van Boxtel, G. J. M., 2000, *Handbook of Psychophysiology. 2nd edition.* Caccioppo, J. T., Tassinary, L. G. & Berntson, G. G. (eds.). New York: Cambridge University Press, p. 507-532 1039 p

Buckolz, E., & Rodgers, R. (1980). The influence of catch trial frequency on simple reaction time. *Acta Psychologica*, *44*(2), 191-200.

Cao, Y., Mahr, A., Castronovo, S., Theune, M., Stahl, C., & Müller, C. A. (2010, February). Local danger warnings for drivers: The effect of modality and level of assistance on driver reaction. In *Proceedings of the 15th international conference on Intelligent user interfaces* (pp. 239-248).

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in cognitive sciences*, *8*(4), 162-169.

Green, D. M., & Von Gierke, S. M. (1984). Visual and auditory choice reaction times. *Acta psychologica*, *55*(3), 231-247.

Hofmann, P., & Rinkenauer, G. (2013). Response preparation in a lane change task. *Ergonomics*, *56*(2), 268-281.

Hofmann, P., Rinkenauer, G., & Gude, D. (2010). Preparing lane changes while driving in a fixed-base simulator: Effects of advance information about direction and amplitude on reaction time and steering kinematics. *Transportation research part F: traffic psychology and behaviour*, *13*(4), 255-268.

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive psychology*, *12*(4), 523-553.

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & cognition*, *7*(3), 166-174.

Lu, C. H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic bulletin & review*, *2*, 174-207.

Lundkvist, A., & Nykänen, A. (2016). Response times for visual, auditory and vibrotactile directional cues in driver assistance systems. *SAE International journal of transportation safety*, *4*(1), 8-14.

Moore, K. S. (2010). *Involuntary orienting of attention to irrelevant task goals* (Doctoral dissertation, University of Michigan).

Müller-Gethmann, H., Rinkenauer, G., Stahl, J., & Ulrich, R. (2000). Preparation of response force and movement direction: Onset effects on the lateralized readiness potential. *Psychophysiology*, *37*(4), 507-514.

Ngo M. and Spence C., "Crossmodal Facilitation of Masked Visual Target Identification," Attention, Perception & Psychophysics 72(7):1938-1947, 2010.

Requin, J., Brener, J., & Ring, C. (1991). Preparation for action In: Jennings JR, Coles MG, editors. Handbook of Cognitive Psychophysiology: Central and autonomic nervous system approaches.

Rosenbaum, D. A. (1980). Human movement initiation: specification of arm, direction, and extent. *Journal of Experimental Psychology: General*, *109*(4), 444.

Sharma, N. R., Kushvah, J. P., & Rinkenauer, G. (2022). Attention and action preparation during lane change maneuvers: The role of irrelevant information. *Human Factors in Transportation*, *60*(60).

Sun, Y. (2016). *Multisensory cue congruency in lane change test*. Michigan Technological University.

Ulrich, R., Schröter, H., Leuthold, H., & Birngruber, T. (2015). Automatic and controlled stimulus processing in conflict tasks: Superimposed diffusion processes and delta functions. *Cognitive psychology*, *78*, 148-174.

# Impact of unrelated driving thoughts on visual processing during manual driving

Alexandra Fort [1]

[1] Laboratory of Ergonomics and Cognitive Science for Transport (LESCOT), Université Gustave Eiffel, Cité des Mobilités, 25 av. François Mitterrand, 69500 Bron, France
[*]alexandra.fort@univ-eiffel.fr

**Abstract: Cognitive distraction due to unrelated driving thoughts (UDT) is known to be a contributory factor to road crashes. This presentation aims to present and compare the results of three complementary studies aiming to better understand the impact of different UDT on manual driving using electrophysiological data combined with behavioral data. The data demonstrate the practical application of Event-Related Potentials (ERPs) to disentangle the impacts of different UDT in the context of driving.**

## 1. Introduction

Driving is a complex and dynamic activity requiring the processing of a large amount of information to assess the situation, anticipate its evolution, and make appropriate decisions, all within a significant time constraint. In this context, the slightest lapse of attention can impair driving performance with deleterious repercussions on safety. Data show that at least 1/3 of road accidents would be caused by a lack of driver attention (Qin et al., 2019).

These attentional deficiencies may come from vigilance issues or activities involving visuo-motor distractions, which have directly observable behavioral markers. However, they can also be induced by internal distractions due to thoughts unrelated to the driving task (UDT), known as mind-wandering episodes. With no directly observable markers, these thoughts are difficult to detect while driving. However, they may occur frequently and they have been identified as a contributory factor in road accidents (Berthié et al., 2015). They have also been associated with an increased likelihood of being at fault in road accidents (Galéra et al., 2012; Lagarde et al., 2004). Thus, although the drivers' gaze is focused on the road, it seems that they are less able to process information and react appropriately when certain thoughts interfere with their driving activity.

If mind-wandering is often used as a generic term referring to a large spectrum of thoughts, these thoughts may be classified further. Three main types may be distinguished: 1) spontaneous thoughts, 2) obsessive thoughts and ruminations, and 3) goal-directed thoughts (Fox & Christoff, 2018). Spontaneous thoughts are generally unfocused with transitions and associations that can be perceived as random. The obsessive thoughts and ruminations may be defined as recurrent, persistent, and unintentional thoughts that persist in a continuous loop in a person's mind. The ruminations are considered distinct from spontaneous thoughts because of their strong focus on precise content and are often linked to a negative mood like sadness or anger, with very limited vagrancy of thoughts. Goal-directed thoughts, or serious thoughts, can also be separated from so-called spontaneous thoughts as they respond to a conscious need from the individual and are therefore perceived as intentional.

In this paper, a synthesis of 3 individual studies, realized in our laboratory is presented, each one aiming to better understand the impact of one of these 3 kinds of UDT on driving behavior and information processing using electrophysiological data. These studies have been firstly published separately (study 1: Bueno et al., 2013; Study 2: Techer et al., 2017; Study 3: Pepin et al., 2020).

## 2. Method

### *2.1 Participants*

Twelve healthy right-handed participants (age: 30.6±3.8) took part in the first experiment (Bueno et al. 2013). Twenty-four participants (age: 32.3 ±5.5) were involved in the second one (Techer et al. 2017) and thirty volunteers (age: 26.88 ± 4.1), in the third (Pepin et al.2020). They reported normal or corrected to normal vision, no neurologic disease, and no medical treatment. The research protocols followed The Code of Ethics of the World Medical Association.

### *2.2 Driving scenario*

Participants were required to drive on a simplified simulator composed of a 24-inch screen, an adjustable car seat, a steering wheel, and three pedals. Participants were required to drive on a straight country road and to follow a motorcycle that braked regularly along the scenario. They had to remove their right foot from the accelerator pedal when they perceived the previous vehicle's brake light (target). Foggy conditions with no traffic were chosen specifically to reduce saccadic movements and to justify, to a certain extent, the frequent decelerations of the motorcycle presented pseudo-randomly from 4 to 12 sec.

In study 1, an auditory warning could forewarn participants that the motorcycle was going to brake soon, in half of the scenarios. In study 2, the warning signal was always present. In study 3, no warning signal was used.

### *2.3 Unrelated Driving Thoughts (UDT)*

In study 1, UDT consisted of serious thoughts induced by a problem-solving task: a set of three words with apparently no links between them was given orally (ex: HAPPY/BLUE/HONEY) and the participant had to find a fourth word linked to each one of the three words (solution: MOON).

In study 2, an anger state was induced before the driving scenario to generate ruminations.

In study 3, UDT consisted of spontaneous thoughts reported by participants at the end of 12 2-min driving scenarios using a continuous scale from 0 to 100, with 0: "I was not focused on driving at all", and 100: "I was perfectly focused on driving".

*2.4 Measures and analysis*

The electrophysiological data were recorded using Biosemi ActiveTwo system (34 electrodes). Evoked potentials related to the visual target (leading vehicle's brake light) were extracted from EEG separately for the different conditions (with or without warning) and according to the attentional state of the participant (control or UDT). Two components were analyzed: the visual N1 and the P3. Anova and t-test were performed to compare the amplitude of each ERP component and the RT to the visual target according to the condition (control/UDT).

## 3. Results

*3.1 Study 1: Goal-directed thoughts*

A significant main effect of UDT ($p<.001$) indicates that participants reacted significantly faster in control condition than in UDT condition, with or without warning.

A main effect was obtained on the amplitude of the N1 ($p<.001$) and of the P3 components ($p=.002$). The amplitude of N1 and P3 was lower in the presence of UDT (−9.40 µV; 8.07 µV, respectively) than in the control condition (−12.56 µV; 12.25 µV, respectively) with or without warning.

*3.2 Study 2: Ruminations due to anger*

Only an impact at the level of the N1 component was observed: visual N1 amplitude ($p<.05$) was smaller during Anger (-7.16 µV) than in Control condition (-8.21 µV). No RT effect was observed.

*3.3 Study 3: Spontaneous thoughts*

Reaction times were longer for the UDT (532ms±110) than for the control condition (471ms ±90).

A significant effect of the UDT was observed on visual N1 amplitude, which was smaller (-4.08 µV) than in control condition (-4.77 µV) ($p < .05$) greater in the AD condition than for the MW The analysis also revealed a significant effect of the UDT on P3 amplitude, which was smaller (3.21 µV) than in the control condition (4.07 µV) ($p<.01$).

## 4. Discussion

The main results show the same impact of UDT induced by goal-directed thoughts (with or without warning) or by spontaneous thoughts (without warning) on the processing of visually driving relevant information, at the behavioral and neurophysiological levels. Concerning the neurophysiological data, the impact occurs at the perceptual level as well as the cognitive level.

However, the impact of ruminations induced by anger is slightly different with no impact on the RT, and only an impact at the perceptual level at the neurophysiological level.

A reduction of the visual N1 amplitude may be due to a reduction of the attention allocated to the sensorial processing of braking lights through an attenuation of sensorial sensitivity. This phenomenon is commonly reported in electrophysiological studies of MW (for a review see Kam et al., 2022). The absence of impact at the behavioral level in the anger condition may be due to an increase in arousal counteracting the impact of the UDT.

These combined results illustrate how the attentional deficiencies related to unrelated driving thoughts may be due to specific mental states issued from the interactions of several factors and resulting in a certain cognitive workload as well as a certain arousal. The levels of each interact to explain the behavior in the primary task (Dehais et al., 2020). Other factors able to impact mental state (i.e. fatigue, personality traits, environment, primary task demand…) than those observed here should be considered to understand how different aspects of the mental state may differently impact the allocation of attentional resources.

The use of other complementary physiological measures could also complement this knowledge to be able to disentangle the different mental states and their consequences on attentional processes (Nilsson et al., 2022).

## 5. Conclusions

These results contribute to the understanding of the deleterious effects of different kinds of UDT. These data could prove valuable for another burgeoning research field closely tied to automation: the monitoring of mental states.

## 6. Acknowledgments

These works have been conducted within the phD project M. Bueno-Garcia, F. Techer and G. Pépin, with the contribution of researchers C. Jallais, C. Gabaude, J. Navarro, Y. Corson and C. Fabrigoule, and engineers D. Ndiaye, F. Moreau, P. Deleurence, B. Piechnik.

**References**

Berthié, G., Lemercier, C., Paubel, P.-V., Cour, M., Fort, A., Galéra, C., Lagarde, E., Gabaude, C., & Maury, B. (2015). The restless mind while driving : Drivers' thoughts behind the wheel. *Accident; Analysis and Prevention*, *76*, 159-165. https://doi.org/10.1016/j.aap.2015.01.005

Bueno, M., Fort, A., Francois, M., Ndiaye, D., Deleurence, P., & Fabrigoule, C. (2013). Effectiveness of a Forward Collision Warning System in simple and in dual task from an electrophysiological perspective. *Neuroscience Letters*, *541*, 219-223. https://doi.org/10.1016/j.neulet.2013.02.030

Fox, K. C. R., & Christoff, K. (2018). *The Oxford Handbook of Spontaneous Thought : Mind-wandering, Creativity, and Dreaming*. Oxford University Press Inc.

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 (Clinical Research Ed.)*, *345*, e8105.

Kam, J. W. Y., Rahnuma, T., Park, Y. E., & Hart, C. M. (2022). Electrophysiological markers of mind wandering : A systematic review. *NeuroImage*, *258*,

119372.
https://doi.org/10.1016/j.neuroimage.2022.119372
Lagarde, E., Chastang, J.-F., Gueguen, A., Coeuret-Pellicer, M., Chiron, M., & Lafont, S. (2004). Emotional Stress and Traffic Accidents : The Impact of Separation and Divorce. *Epidemiology*, *15*(6), 762. https://doi.org/10.1097/01.ede.0000142140.77194.ad
Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, *69*(3), 220-232. https://doi.org/10.1037/h0048850
Nilsson, E. J., Bärgman, J., Ljung Aust, M., Matthews, G., & Svanberg, B. (2022). Let Complexity Bring Clarity : A Multidimensional Assessment of Cognitive Load Using Physiological Measures. *Frontiers in Neuroergonomics*, *3*, 787295. https://doi.org/10.3389/fnrgo.2022.787295
Pepin, G., Fort, A., Jallais, C., Moreau, F., Ndiaye, D., Navarro, J., & Gabaude, C. (2020). Impact of mind-wandering on visual information processing while driving : An electrophysiological study. *Applied Cognitive Psychology*, *35*. https://doi.org/10.1002/acp.3773
Qin, L., Li, Z. (Richard), Chen, Z., Andi Bill, M. S., & Noyce, D. A. (2019). Understanding driver distractions in fatal crashes : An exploratory empirical analysis. *Journal of Safety Research*, *69*, 23-31. https://doi.org/10.1016/j.jsr.2019.01.004
Techer, F., Jallais, C., Corson, Y., Moreau, F., Ndiaye, D., Piechnick, B., & Fort, A. (2017). Attention and driving performance modulations due to anger state : Contribution of electroencephalographic data. *Neuroscience Letters*, *636*, 134-139. https://doi.org/10.1016/j.neulet.2016.11.011

# Occupant status monitoring: A framework to integrate human factors, technology & policy to realise injury reductions across the safety chain

Michael G. Lenné [1,2*], Michael Fitzharris [2*], Rikard Fredriksson [3,4,5], Adriano Palao [5] & Mark S. Young [6]

[1] Seeing Machines, Melbourne, Victoria, Australia.
[2] Monash University Accident Research Centre, Monash University, Melbourne, Victoria, Australia
[3] Swedish Transport Administration, Skövde, Sweden
[4] Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Göteborg, Sweden
[5] European New Car Assessment Programme (Euro NCAP), Leuven, Belgium
[6] University of Southampton, Southampton, United Kingdom
[*] corresponding.author@second.com

**Abstract: European NCAP protocols for Occupant Status Monitoring (OSM) recognise the potential role for Driver Monitoring Systems (DMS) in crash avoidance. This paper provides a framework that expands the potential application of DMS to crash avoidance and crash protection phases of the Integrated Safety Chain (ISC). We apply the Field of Safe Travel (FoST) to represent the perceived safe pathways through a typical intersection crash scenario. Driver inattention is a common reason these crashes occur, and this can now be assessed in real-time, to at least a large extent, through the increasing deployment of OSM. Research addressing schema theory, driver expectancy and attentional cueing is used to develop HMI principles that should more effectively re-orient driver attention to emerging threats in the environment. Representing through FoST, OSM enables the dynamic representation of a driver's perception of risk. Integrating this with objective risk from the external vehicle sensors presents opportunities for enhanced crash avoidance through tailoring of ADAS strategies around functional adaption and suppression – improving both safety and driver experience.**

## 1. Introduction

The importance of keeping eyes on the road, being alert and engaged in the driving task are well established ingredients of safe driving. Direct occupant monitoring, via interior camera(s), is a promising approach for capturing driver attention and state in real-time. There is demonstrated safety benefit in real-world fleets (Fitzharris et al., 2017), and it is the focus European NCAP protocols for Occupant Status Monitoring (OSM; Fredriksson et al., 2021).

While OSM achieves safety benefits by detecting driver inattention and provides an alert, we believe the perceptual and attentional underpinnings are more nuanced and that the opportunities to deliver stronger outcomes are significant. Here we integrate a long-standing concept, the Field of Safe Travel (FoST: Gibson & Crooks, 1938), with human factors research in attention and cueing and safety research on crash avoidance and crash protection. A framework is proposed that expands the reach for integrated OSM and ADAS approaches to reduce road injury and increase driver acceptance.

## 2. FoST applied to intersection crashes involving driver inattention

The FoST is a means to illustrate a driver's perception of the field of possible paths that a vehicle may take unimpeded and the minimum stopping distance (Gibson & Crooks, 1938). It represents the driver's perceived safety space or buffer, and thus guides their control of the vehicle. The FoST is highly dynamic, includes the driver's predictive estimations on how a traffic scene will evolve, and remains the subject of current research and application (e.g., Kolekar Kolekar, de Winter & Abbink, 2020; Papakostopoulos, Marmaras & Nathanael, 2017).

To illustrate, a driver (A) might not take account of Vehicle E emerging from the side road in their FoST, because they are focused on passing Vehicle B (Fig 1). If the FoST is breached in this way and a crash occurs, is it likely the driver did not perceive (or recognise) the potential threat and/or not interpret the actual risk appropriately (e.g., Beanland et al., 2013; Fitzharris et al., 2022); in other words, their perceived FoST (and consequent driving decisions) did not match the actual situation. Our focus here is on scenarios where a more effective driver response could have prevented the crash (or at least reduced its severity). If viewed as driver inattention, a simplistic view is simply to warn the driver. However, there are nuances here that must be addressed to realise the significant injury reductions anticipated with safety systems including OSM.

The driver schema may not be directing their attention to the right areas in the road environment to detect potential threats (they may not see it), or they may perceive the threat but react too slowly because they did not expect to see this type of behaviour, which is again related to their schema (Neisser, 1976). Schema theory accounts for the role of experience and expectation and how this drives exploration of the environment and what is perceived. The driver may be directing attention off road (Fig 1, scenario 2) and therefore not attending to the appropriate environmental cues, or may ostensibly be visually attentive (Fig 1, scenario 1) but their information processing has resulted in an inappropriate FoST.

The question then becomes: if we know where visual attention is directed in real-time, what opportunities does this afford to enhance safety and the driver experience? Analysis of the safety chain models reveals points of intervention and opportunities for active and passive safety countermeasures.

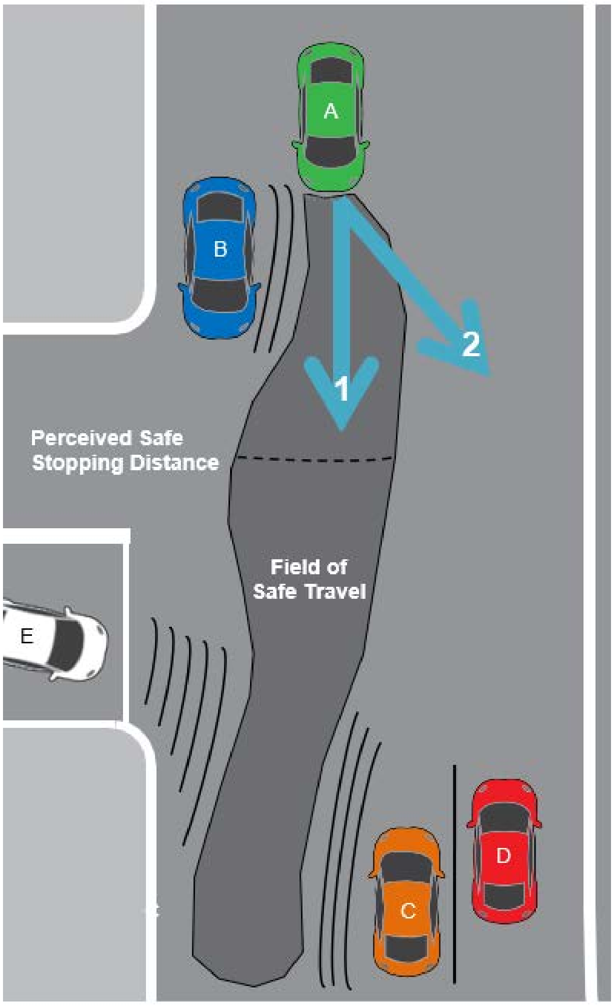



***Fig. 1.** FoST illustration for a driver approaching an intersection. Assumes left-hand drive. Vehicle A represents ego vehicle. Driver attention can be directed straight ahead (scenario 1) or off-road to the right (scenario 2).*

## 3. The Integrated Safety Chain (ISC) and opportunities for intervention

The ISC, initially described by Tingvall (2008) and adopted by others (e.g., Lie 2012; Rizzi 2008), was expanded by Fitzharris et al. (2022a, b). Describing crash phases, it is used to identify countermeasure intervention points, for example, to address impaired driving (Fitzharris et al., 2022b; Lie et al., 2024).

A focus here is on Normal Driving where a vehicle is controlled in a manner that is appropriate to the conditions and results in a safe journey. As the driver moves from Deviations from Normal Driving through to the Critical Situation (Fig 2, phases 3-5), opportunities exist to re-orient driver attention to avoid the crash (Fitzharris et al., 2023). Warning signals that prime the driver and re-orient attention more rapidly will elicit more efficient responses. The use of directional cues within warning systems and, ideally, presentation of warnings in the same spatial location, along with multi-modal cues, are key ingredients to enhance road user expectancy (Ho & Spence, 2005; Ho et al., 2006; Posner, Snyder, & Davidson, 1980; Salmon et al., 2014).

For example, in scenario 1 (Fig 1) a potentially longer warning threshold could be used given the driver's attention is already oriented on-road (see Table 1). An informative visual warning from the instrument cluster or HUD would align with direction of attention, although auditory cues may have a secondary role. An earlier warning could be presented to the driver in scenario 2 given the additional time required to re-orient attention to the forward roadway. A visual alert in the cluster may not be within the driver's field of view therefore an auditory warning, or haptic cues including low-G braking will likely be more effective in capturing attention.

***Table 1.** Key elements of and potential tailoring options for the HMI warning. Reference is made to phases in the ISC (Fig 2) and visual attention scenarios 1 & 2 (Fig 1).*

| | Timing | Modality | Location |
|---|---|---|---|
| Scenario 1 (eyes on road) | Phase 4. Opportunity for delayed warning onset aligned with emerging situation, followed by escalation | Visual | Cluster or HUD |
| Scenario 2 (eyes off road) | Earlier warning in the emerging situation (phase 3), steep escalation in warning intensity (phase 4) | Auditory or haptic | Use spatial cues for faster re-orientation of attention to the road |

## 4. Discussion

Driver schema direct attention to objects perceived as salient in the environment, mapped by the FoST. With driver attention able to be measured in real-time through OSM and vehicle sensors mapping the external environment in real-time, bringing these two together enables driver-vehicle collaboration to really evolve through smarter ADAS that affords new functional adaptation and suppression strategies. As OSM approaches typically capture visual attention and not cognitive distractions, at this stage, the importance of the core ADAS suite of external sensing technologies remains; indeed, there is an opportunity to build upon them for safer travel.

By means of example, FCW is typically developed under the assumption that the driver is (visually) distracted, hence embedding driver state into the function could be highly beneficial. If a driver is looking left, and a vehicle approaches from the right (Fig 1, scenario 2), a higher sensitivity could be used that incorporates directional cues to more rapidly re-direct attention to the right (functional adaptation). Conversely, if a driver is already focussed on the forward roadway, or there are no threats and the FoST is large, there may be opportunities to delay or suppress warnings in some circumstances (Fig 1, scenario 1). Theoretically a greater proportion of critical warnings would be issued when drivers need them and fewer when they do not. Supporting a good driver experience with respect to expectations of OSM performance, including false detection rates, is critical to achieve the driver acceptance and engagement necessary to deliver the desired safety outcomes.

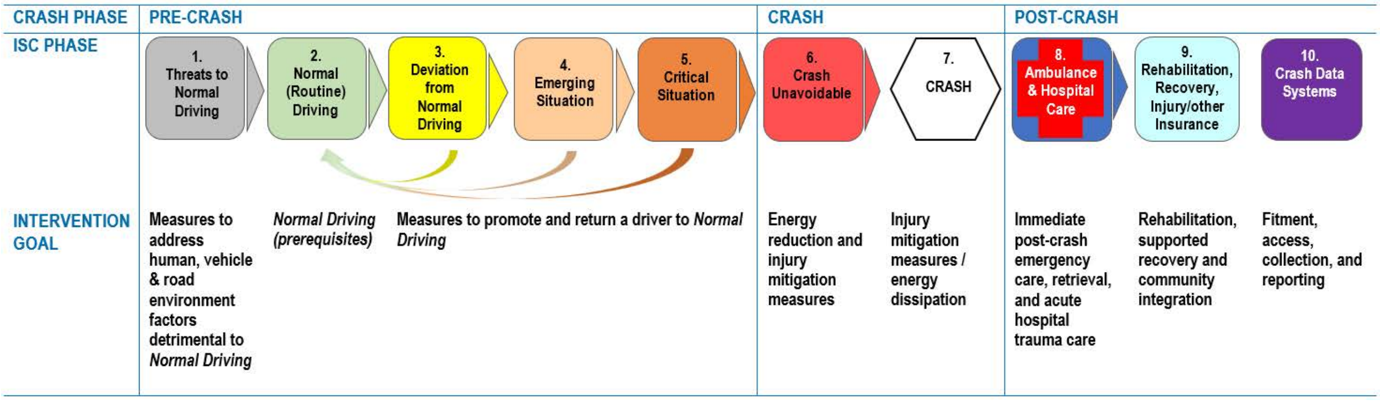



***Fig. 2.*** *Integrated safety chain (Fitzharris et al., 2022a)*

Examining driver behaviour in the ways outlined here confirms current policy decisions by Euro NCAP. Greater reward is given to OSM implementations that seek to avoid a crash through provision of a safety intervention, rather than a warning only. This provides intervention opportunities as the deviation from normal driving becomes critical. OSM strategies that leverage functional adaptation and suppression strategies are key here.

Further opportunities to strengthen crash avoidance through management of other behaviours that draw attention off road (e.g., passenger interactions, being out of position, etc) exist. Activation of reversible passive protection, for example electric seatbelt pretensioners with increasing pull force levels to increase safety whilst acting as a haptic warning is another promising direction. Knowledge of occupant position and size from OSM systems could also be used in future to adapt the deployment of airbag systems.

## 5. Conclusions

Integrating driver perception of risk through their FoST with objective risk presents a significant road safety opportunity, even more so with assisted driving where disparity between perceived and actual risk may be greater (de Winter et al. 2023). OSM approaches enable real-time assessment of driver attention that can be mapped against the external threats mapped by vehicle sensors, the outputs of which can be used to tailor ADAS strategies around functional adaption and suppression. Leveraging principles of attentional capture and expectancy provide new opportunities to increase driver acceptance and realise greater safety gains.

## 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 In-depth Study. *Accident Analysis and Prevention, 54,* 99-107.

De Winter, J., Hoogmoed, J., Stapel, J., Dodou, D. & Bazilinskyy, P. (2023). Predicting perceived risk of traffic scenes using computer vision. *Transportation Research Part F: Psychology & Behaviour, 93,* 235-247.

Fitzharris, M., Corben, B., Lenné, M., Liu, S., Peiris, A., Pok Arundell, T., Stephens, A.N., Bowman, D., Morris, A., Tingvall, C. (2022a*). Understanding contributing factors for serious injury crashes using crash chain analysis*, ECIS Report 3 (MUARC Report 345). doi:10.26180/23305382

Fitzharris, M., Pok Arundell, T., Corben, B., Lenné, M., Liu, S., Peiris, A., Stephens, A.N., Bowman, D., Morris, A., Tingvall, C. (2022b). *Identification of countermeasures to address serious injury crashes*, ECIS Report 4 (MUARC Report 346). doi:10.26180/23305388

Fitzharris, M., Lenné, M., Liu, S., Pok Arundell, T., Peiris, A., Tingvall, C., Corben, B., Bowman, D., Morris, A. (2023). Application of a modified Integrated Safety Chain using in-depth crash data to identify factors associated with serious injury crashes: a method to prioritise currently available active safety systems and to identify new opportunities to advance vehicle safety. *The 27th ESV Conference Proceedings, The 27th International Technical Conference on the Enhanced Safety of Vehicles (ESV)* Yokohama, Japan, April 3-6, 2023. NHTSA. doi:10.26180/23365655

Fitzharris, M., Liu, S., Stephens, A. N. & Lenné, M. G. (2017) The relative importance of real-time in-cab and external feedback in managing fatigue in real-world commercial transport operations. *Traffic Inj. Prev. 18,* S71–S78.

Fredriksson, R., Lenné, M. G., van Montfort, S., & Grover, C. (2021). European NCAP Program Developments to Address Driver Distraction, Drowsiness and Sudden Sickness. *Frontiers in Neuroergonomics, 2.* https://www.frontiersin.org/articles/10.3389/fnrgo.2021.786674/full

Gibson, J. J. & Crooks, L. E. (1938). A theoretical field-analysis of automobile driving. *American Journal of Psychology, 51,* 453-471.

Kolekar, S., de Winter, J. & Abbink, D. (2020). Which parts of the road guide obstacle avoidance? Quantifying the driver's risk field. *Applied Ergonomics, 89,* 103196.

Lie A (2012). *Managing traffic safety: an approach to the evaluation of new vehicle systems*. Stockholm, Sweden: Karolinska Institute.

Lie, A., Tingvall, C., Michael, J. P., Fell, J. C. & Bella Dinh-Zarr, T. (2024). Vision Zero and Impaired Driving: Near and Longer-Term Opportunities for Preventing Death and Injuries. *Accident Analysis & Prevention, 194,* 107344.
Ljung Aust, M. (2023). Key human factors development principles for DMS enhanced collision avoidance system development. Paper presented at The 27th Enhanced Safety of Vehicles, Yokohama, Japan.

Neisser, U. (1976). *Cognition and Reality: Principles and Implications of Cognitive Psychology*. San Francisco: Freeman.

Palao, A., Fredriksson, R., & Lenné, M. (2023). Euro NCAP's current and future in-cabin monitoring systems assessment. *In 27th International Technical Conference on the Enhanced Safety of Vehicles (ESV),* National Highway Traffic Safety Administration (No. 23-0286).

Papakostopoulos, V., Marmaras, N. & Nathanael, D. (2017). The "field of safe travel" revisited: interpreting driving behaviour performance through a holistic approach. *Transport Reviews, 37*, 695-714.

Salmon, P. M., Lenné, M. G., Walker, G. H., Stanton, N. A. & Filtness, A. (2014). Exploring schema-driven differences in situation awareness between road users: an on-road study of driver, cyclist and motorcyclist situation awareness, *Ergonomics, 57,* 191-209.

Rizzi M (2016). Towards a safe system approach to prevent health loss among motorcyclists. *PhD thesis: Chalmers University of Technology*: Gothenburg.

Tingvall C. (2008). Distraction from the view of governmental policy making subsection of the chapter: Government and Industry Perspectives on Driver Distraction. In: Regan M, Lee J, Young K, (eds). *Driver distraction: Theory, effects, and mitigation.* CRC Press; Melbourne.

# Detection of Mind Wandering during Simulated Delegated Driving: Influence on Physiological Measurements

Christophe Jallais [1], Florence Mazeres[1], Yannick Lavril[1], Hélène Tattegrain[1], Daniel Ndiaye[1], Bertrand Richard[1], Alexandra Fort[1]

[1] Laboratory of Ergonomics and Cognitive Science for Transport (LESCOT), Université Gustave Eiffel, Cité des Mobilités, 25 av. François Mitterrand, 69500 Bron, France
[*]christophe.jallais@univ-eiffel.fr

Abstract: **This study investigates the impact of Mind Wandering (MW) on physiological measurements during simulated delegated driving. Participants were asked to supervise the driving environment during the 20-minute driving delegation session. Every 60 to 120 seconds, participants had to report whether their attention was focused on the supervision task or on internal thoughts (MW). They declared their focus 11 times, with the minute before each response categorized as "mind-wandering" or "focus on the task.". Results demonstrate a significant effect of MW on respiratory and electrodermal indicators, with individuals in the MW group exhibiting reduced sighing and phasic electrodermal responses compared to those in the task-focused (monitoring the driving environment) group. These findings suggest a decrease in task engagement during MW, consistent with literature associating MW with tasks requiring minimal engagement. Findings highlight the importance of considering various physiological indicators to detect mind-wandering episodes. However, the reliability of these indicators in manual driving contexts warrants further investigation.**

## 1. Introduction

According to Singh (2015), 90% of road accidents are due to human error, which can be caused by degraded states of driver attention. While research has uncovered a specific signature of visual distraction (e.g., glances off the road), this is not the case for inattention, which is said to be the 3rd factor in freeway accidents responsible for 13% of fatal accidents in France (ONISR, 2022). Furthermore, according to Galera et al. (2012), wandering thoughts increase the risk of being responsible for or involved in an accident (Lagarde et al., 2004), or even engaging in risky driving (Lemercier et al., 2014; Yanko & Spalek, 2014). Mind wandering (MW) is a spontaneous state corresponding to a shift of thoughts related to a task at hand (e.g., driving) to self-generated thoughts and feelings (Smallwood & Schooler, 2015). Additionally, people who report a high propensity for MW also report more infractions and driving failures (Burdett et al., 2016). The question of detecting these states while driving (manual or delegated) therefore becomes crucial.

In manual driving, the impact of MW has already been assessed using behavioral measures (lane position, speed) and ocular measures (blink, eye movement). To our knowledge, few studies have investigated the impact of MW on the autonomic nervous system and found its specific signature (Albert et al., 2022; Bortolla and Maffei, 2022; Gontier, 2017). However, autonomic system measurements such as heart rate and electrodermal conductance have demonstrated their relevance for detecting periods of boredom (Perkins, 1981), or for determining pilot vigilance in real-time (Boucsein et al., 2007). Building on this, MW has been associated with simple tasks requiring few resources (Randall et al., 2019), this type of task being associated with low autonomic system activity (Obrist, 1981). Given the results of previous studies, we can assume that thought wandering may be associated with low autonomic nervous system activity.

This study aims to detect the presence of MW episodes by identifying their effects on physiological indicators. A main effect of MW on physiological parameters during autonomous driving reflecting low resource mobilization and engagement in the supervisory task is expected (e.g. reduced heart rate, tonic skin conductance level, respiratory rate).

## 2. Method

### *2.1 Participants*

Forty-three participants (mean age = 29.09 years, SD = 11.61, 27 women) took part in this study. They had to have at least three years' driving experience, be right-handed, and be aged between 21 and 45.

### *2.2 Materials and Measurements*

The simulator comprised a Peugeot 308 cabin surrounded by 8 screens (220 cm high × 165 cm wide), providing a horizontal field of view of approximately 280° and a vertical field of view of around 40°.

Physiological data, including cardiac, respiratory, and electrodermal signals, were collected and synchronized using RTMaps software. The mean heart rate (HR) and the heart rate variability indicator (RMSSD) were calculated. Additionally, respiratory rate (RR), respiratory volume, number of sighs, and respiratory variability (BRV) were computed by detecting peak inspirations on the filtered respiratory signal. Electrodermal response characteristics were calculated using MATLAB Toolbox Ledalab.

### *2.3 Procedure*

Before commencing the experiment, the experimenter equipped the participants with the respective electrodes for each tool (ECG, EDA, and EEG). Participants then engaged in delegated driving mode on a driving simulator. The driving session lasted 20 minutes, conducive to mind wandering (see Baldwin et al., 2017), during which participants were instructed to supervise the driving environment. Throughout the session, participants were prompted to report every 60 to 120 seconds whether their attention was focused on the supervision task or not, facilitated by an auditory signal indicating the presentation of a questionnaire on a touch-sensitive tablet attached to the vehicle cockpit. They were required to declare their attentional state 11 times during the experiment. The minute preceding each response could be labeled as either "mind-wandering" or "focus on the task (supervising the driving environment)".

*2.4 Analysis*

We calculated the averages for the five-minute rest for all physiological parameters. To account for inter-individual variability, we normalized all physiological data by subtracting the mean of the rest responses from the 60 seconds before the onset of a questionnaire (Llabre et al., 1991). This is a standard approach for analyzing physiological responses induced by a task or state change (e.g., Mazeres et al., 2019). To test our hypotheses regarding the impact of MW on physiological responses, we compared the periods preceding the responses of the two groups "MW" vs. "focus on task" as a function of questionnaire numbers.

**3. Results**

The means and standard errors of the physiological reactivity scores according to the drivers' attentional class (on task vs. mind wandering) are presented in Table 1. An ANOVA with thought type (mind wandering or not) and questionnaire number (1 to 11) as between-subjects factors was performed for all indicators. Planned comparisons were made when the interaction effect was significant or trend.

***Table 1*** Physiological score means and standard errors

| | **Focus on Task** | **MW** |
|---|---|---|
| | **Cardiovascular reactivity** | |
| **RMSSD** | -.002 (.016) | -.003 (.015) |
| **HR** | -.92 (4.49) | -.69 (4.90) |
| | **Respiratory reactivity** | |
| **BR** | 1.87 (2.45) | 1.54 (2.31) |
| **Volume** | -.11 (.33) | -.14 (.36) |
| **Number of sighs** | -.08 (.55) | -.34(.76) |
| **BRV** | -134.80 (664.68) | -97.48 (625.51) |
| | **Electrodermal reactivity** | |
| **Amplitude** | .025 (.14) | .002 (.17) |
| **Number of SCR** | -4.21 (6.83) | -6.02 (7.57) |
| **Tonic** | .116 (.494) | .027 (.57) |

*Note. RMSSD and BRV are indicated in milliseconds, HR and BR are indicated in beats and breaths per minute respectively, and amplitude and tonic are indicated in microsiemens*

*3.1 Cardiovascular reactivity*

RMSSD reactivity analyses revealed that the interaction effect did not reach significance, $F(10, 451) = 1.047$, $p = .403$, $\eta 2p = .023$. The main effect of questionnaire number and the effect of thought type were not significant, $F(10, 451)= .467$, $p = .911$, $\eta p2= .010$ and $F(1, 451)= .011$, $p = .916$, $\eta p2 < .001$.

HR reactivity analyses revealed that the interaction effect was not observed, $F(10, 451)= .910$, $p = .524$, $\eta 2p = .020$. The main effect of questionnaire number and the effect of thought type were not significant, $F(10, 451)= .659$, $p = .762$, $\eta p2=.014$ and $F(1, 451)= .221$, $p = .639$, $\eta p2 < .001$.

*3.2 Respiratory reactivity*

Analyses of respiratory rate (RR) reactivity did not reveal an interaction effect, $F(10, 451)= .916$, $p = .518$ $\eta 2p = .020$. The main effect of questionnaire number and the effect of thought type were not significant, $F(10, 451)= .410$, $p = .942$, $\eta p2=.009$ and $F(1, 451)= 2.847$, $p = .092$, $\eta p2=.006$.

Analyses of the number of sighs did not show an interaction effect, $F(10, 451) = 0.210$, $p = 0.995$ $\eta 2p = 0.005$. The main effect of the questionnaire number was not significant, $F(10, 451)= .861$, $p = .570$, $\eta p2=.019$ but the main effect of thought type was observed, $F(1, 451)= 12.871$, $p < .001$, $\eta p2=.028$. When individuals had wandering thoughts, they had a low number of sighs (see Fig.1).

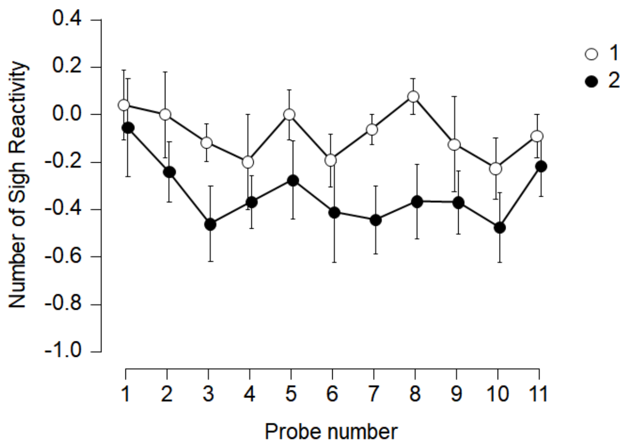



**Fig. 1.** Means and standard errors of the reactivity of the number of sighs for each questionnaire according to the type of thoughts (1: on-task or 2: MW).

BRV reactivity analyses revealed that the interaction effect was a trend, $F(10, 451) = 1.714$, $p = 0.075$ $\eta 2p = 0.037$. The main effect of questionnaire number and the effect of thought type were not significant, $F(10, 451)= 1.521$, $p = .129$, $\eta p2=.033$ and $F(1, 451)= .778$, $p = .378$, $\eta p2=.002$.

*3.3 Electrodermal reactivity*

Analyses of tonic component response reactivity did not reveal an interaction effect, $F(10, 451) = .421$, $p = .936$ $\eta 2p = .009$. The main effect of questionnaire number and the effect of thought type were not significant, $F(10, 451)=.578$, $p = .832$, $\eta p2=.013$ and $F(1, 451)= 1.759$, $p = .185$, $\eta p2=.004$. All effects were also non-significant for electrodermal response amplitude, $Fs < .2,568$ and $p > .110$.

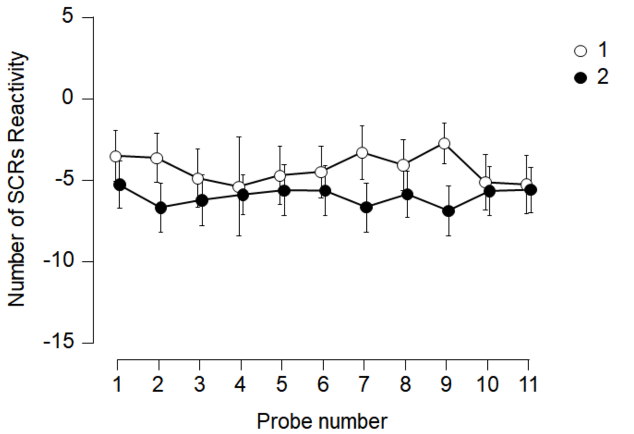



**Fig. 2.** Means and standard errors of the reactivity of the number of phasic electrodermal responses for each questionnaire according to the type of thoughts (1: on-task or 2: MW).

Analyses of SCR numbers showed that the interaction effect did not reach significance, $F(10, 451) = 0.273$, $p = 0.987$ $\eta 2p = 0.006$. The main effect of questionnaire number was not significant, $F(10, 451)=.084$, $p = 1.000$, $\eta p2=.002$. However, the effect of thought type was significant $F(1, 451)= 5.123$, $p = .024$, $\eta p2=.011$. When individuals had wandering thoughts, they had a low number of SCRs irrespective of probe time (see Fig. 2).

## 4. Discussion

The results revealed a significant effect of MW on respiratory and electrodermal indicators: the "MW" condition sighed less and exhibited fewer phasic electrodermal responses compared to those in the "on task" condition.

These findings suggest diminished task engagement when individuals experience MW. In the literature, MW is associated with tasks requiring minimal engagement (Randall et al., 2019). In the context of autonomous driving, individuals could allocate fewer resources to supervision and engage in thoughts unrelated to driving. However, there was no significant effect on cardiovascular activity. It may be worthwhile to explore other, potentially more sensitive, measures of cardiac activity for a more precise assessment of attentional engagement. For instance, the pre-ejection period (PEP) has been identified as a more direct and non-invasive measure commonly used in engagement studies (Mazeres et al., 2021).

Some studies have examined brain activity to detect MW, reflected by perceptual decoupling (e.g., Baldwin et al., 2017; Pepin et al., 2021). Comparing these different indicators and investigating the connection between energetic disengagement and attentional decoupling could be an interesting avenue for future research.

The results of this study reveal a specific electrodermal and respiratory signature of wandering thoughts, indicating decreased engagement in the task at hand. While participants drove in delegated mode, the reliability of these indicators in manual driving scenarios requires verification. Nonetheless, it is essential to consider these various indicators to detect degraded attentional states, such as wandering thoughts, during driving, utilizing portable tools that continuously monitor drivers' physiological responses (e.g., smartwatches).

## 5. Acknowledgments

The authors would like to express their gratitude to the participants of this study. This research was funded by BPI France as part of the SERENITE project (CORAM 2021).

## References

Albert, D. A., Claude Ouimet, M., & Brown, T. G. (2022). Negative mood mind wandering and unsafe driving in young male drivers. Accident Analysis & Prevention, 178, 106867. https://doi.org/10.1016/j.aap.2022.106867

Baldwin, C. L., Roberts, D. M., Barragan, D., Lee, J. D., Lerner, N., & Higgins, J. S. (2017). Detecting and quantifying mind wandering during simulated driving. Frontiers in Human Neuroscience, 11(406). https://doi.org/10.3389/fnhum.2017.00406

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1993). Cardiac psychophysiology and autonomic space in humans: Empirical perspectives and conceptual implications. Psychological Bulletin, 114(2), 296–322. https://doi.org/10.1037/0033-2909.114.2.296

Boucsein. (2012). Publication recommendations for electrodermal measurements. Psychophysiology, 49(8), 1017 1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x

Boucsein, W., Haarmann, A., & Schaefer, F. (2007). Combining skin conductance and heart rate variability for adaptive automation during simulated IFR flight. In D. Harris (Éd.), Engineering Psychology and Cognitive Ergonomics (p. 639 647). Springer. https://doi.org/10.1007/978-3-540-73331-7_70

Bortolla, R., Galli, M., Spada, G. E., & Maffei, C. (2022). Mindfulness effects on mind wandering and autonomic balance. Applied Psychophysiology and Biofeedback, 47(1), 53 64. https://doi.org/10.1007/s10484-021-09527-x

Burdett, B. R. D., Charlton, S. G., & Starkey, N. J. (2016). Not all minds wander equally: The influence of traits, states and road environment factors on self-reported mind wandering during everyday driving. Accident Analysis & Prevention, 95, 1-7. https://doi.org/10.1016/j.aap.2016.06.012

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (2007). Handbook of psychophysiology. Cambridge university press.

Carriere, J. S. A., Cheyne, J. A., & Smilek, D. (2008). Everyday attention lapses and memory failures: The affective consequences of mindlessness. Consciousness and Cognition, 17(3), 835-847. https://doi.org/10.1016/j.concog.2007.04.008

Framorando, D., & Gendolla, G. H. E. (2018). The effect of negative implicit affect, prime visibility, and gender on effort-related cardiac response. Adaptive Human Behavior and Physiology, 4(4), 354 363. https://doi.org/10.1007/s40750-018-0097-0

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, 345, e8105. https://doi.org/10.1136/bmj.e8105

Gontier, C. (2017). How to prevent mind-wandering during an EVA ? Presentation of a mind-wandering detection method using ECG technology in a Mars-analog environment. Acta Astronautica, 140, 105 112. https://doi.org/10.1016/j.actaastro.2017.08.008

Lagarde, E., Chastang, J.-F., Gueguen, A., Coeuret-Pellicer, M., Chiron, M., & Lafont, S. (2004). Emotional Stress and traffic accidents:The impact of separation and divorce. Epidemiology, 15(6), 762-766.

Llabre, M. M., Spitzer, S. B., Saab, P. G., Ironson, G. H., & Schneiderman, N. (1991). The reliability and specificity of delta versus residualized change as measures of cardiovascular reactivity to behavioral challenges. Psychophysiology, 28(6), 701–711. https://doi.org/10.1111/j.1469-8986.1991.tb01017.x

Lemercier, C., Pêcher, C., Berthié, G., Valéry, B., Vidal, V., Paubel, P.-V., Cour, M., Fort, A., Galéra, C., Gabaude, C., Lagarde, E., & Maury, B. (2014). Inattention behind the wheel : How factual internal thoughts impact attentional control while driving. Safety Science, 62, 279 285. https://doi.org/10.1016/j.ssci.2013.08.011

Mazeres, F., Brinkmann, K., & Richter, M. (2019). Implicit achievement motive limits the impact of task difficulty on effort-related cardiovascular response. Journal of Research in Personality, 82. https://doi.org/10.1016/j.jrp.2019.06.012

Mazeres, F., Brinkmann, K., & Richter, M. (2021). Motivated but not engaged: The implicit achievement motive requires difficult or unclear task difficulty conditions to exert an impact on effort. Journal of Research in Personality, 94. https://doi.org/10.1016/j.jrp.2021.104145

Maillet, D., Beaty, R. E., Jordano, M. L., Touron, D. R., Adnan, A., Silvia, P. J., Kwapil, T. R., Turner, G. R., Spreng, R. N., & Kane, M. J. (2018). Age-related differences in mind-wandering in daily life. Psychology and Aging, 33(4), 643 653. https://doi.org/10.1037/pag0000260

Mueller, A. S., Reagan, I. J., & Cicchino, J. B. (2021). Addressing driver disengagement and proper system use: Human factors recommendations for Level 2 driving automation design. Journal of Cognitive Engineering and Decision Making, 15(1), 3 27. https://doi.org/10.1177/1555343420983126

Obrist, P. A. (1981). Cardiovascular psychophysiology : A perspective. Plenum Press.

Observatoire National Interministériel de la Sécurité Routière [ONISR]. (2022). La sécurité routière en France - Bilan de l'accidentalité de l'année 2016.

Pepin, G., Fort, A., Jallais, C., Moreau, F., Ndiaye, D., Navarro, J., & Gabaude, C. (2021). Impact of mind-wandering on visual information processing while driving : An electrophysiological study. Applied Cognitive Psychology, 35(2), 508 516. https://doi.org/10.1002/acp.3773

Perkins, R. E. (1981). The nature and origins of boredom [Keele University].

Poerio, G. L., Totterdell, P., & Miles, E. (2013). Mind-wandering and negative mood : Does one thing really lead to another? Consciousness and Cognition, 22(4), 1412 1421. https://doi.org/10.1016/j.concog.2013.09.012

Randall, J. G., Beier, M. E., & Villado, A. J. (2019). Multiple routes to mind wandering : Predicting mind wandering with resource theories. Consciousness and Cognition, 67, 26 43. https://doi.org/10.1016/j.concog.2018.11.006

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), 287 300. https://doi.org/10.1037/a0034386

Singh, S. (2015). Critical reasons for crashes investigated in the national motor vehicle crash causation survey. National Highway Traf ic Safety Administration, Tech. Rep. DOT HS 812 115, February

Smallwood, J., & O'Connor, R. C. (2011). Imprisoned by the past : Unhappy moods lead to a retrospective bias to mind wandering. Cognition and Emotion, 25(8), 1481 1490. https://doi.org/10.1080/02699931.2010.545263

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

Tamir, M. (2016). Why do people regulate their emotions? A taxonomy of motives in emotion regulation. Personality and Social Psychology Review, 20(3), 199 222. https://doi.org/10.1177/1088868315586325

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. Scientific Reports, 7(1), Art. 1. https://doi.org/10.1038/s41598-017-14849-0

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

Young, M. S., & Stanton, N. A. (2002). Attention and automation : New perspectives on mental underload and performance. Theoretical Issues in Ergonomics Science, 3(2), 178 194. https://doi.org/10.1080/14639220210123789

# Sadness Detection in Autonomous Driving: A Physiological Approach

Christophe Jallais [1]

[1] Laboratory of Ergonomics and Cognitive Science for Transport (LESCOT), Université Gustave Eiffel, Cité des Mobilités, 25 av. François Mitterrand, 69500 Bron, France
*christophe.jallais@univ-eiffel.fr

**Abstract:** The driver's attentional state plays a crucial role in road safety, as it directly contributes to accidents. Two studies explored the feasibility of detecting sadness using physiological measures such as ECG, respiration, and EDA. The results indicate that accurately classifying sadness during autonomous driving is possible. However, the data were collected using laboratory systems, necessitating technological advancements to acquire similar data in real-world motor vehicles.

## 1. Introduction

Episodes of inattention are frequently observed during manual driving, diverting the driver's attention from the primary task of driving. Studies have highlighted an increased risk of being involved in at-fault accidents when drivers experience inattention due to rumination linked to sadness (Lagarde et al., 2004). Hence, it is imperative to monitor, identify, and address drivers' internal states, particularly their emotions.

Existing literature indicates that negative emotions, such as sadness, can lead to aggressive behavior, impaired driving performance, and risky driving (Garrity & Demick, 2001; Jallais et al., 2014; Stephens & Groeger, 2011). Despite the recognized impact of emotions on driving, especially the influence of sadness-related ruminations, there is a notable absence of studies attempting to identify and mitigate emotional states using physiological parameters in autonomous driving scenarios.

Distractive thoughts can manifest either voluntarily, with control and purpose, such as when contemplating future events or problem-solving, or spontaneously, without control or direction, like daydreaming or ruminative thoughts associated with negative emotions (anger, sadness, anxiety, etc.). These thoughts divert attention from driving-related tasks, competing for space in working memory and resulting in perceptual decoupling, where attention shifts away from external stimuli (Christoff et al., 2016; Smallwood & Schooler, 2015).

Capturing distractive thoughts, such as ruminations, poses challenges due to their lack of observable events. To investigate them effectively, a methodological triangulation of subjective, behavioral, and physiological data is necessary (Gruberger et al., 2011). Numerous studies have explored the physiological responses linked to emotional states using metrics like heart rate, recorded via electrocardiogram (ECG) (Mesken et al., 2007; Stemmler et al., 2007).

To address this gap, we conducted two simulator-based studies to investigate the physiological indicators relevant to detecting negative emotional states, specifically sadness, during driving. We aimed to identify which physiological indicators are most sensitive for classifying the state of sadness in the context of automobile driving. Therefore, this paper presents the findings from two studies to demonstrate the effectiveness of physiological indicators in classifying the state of sadness, as well as identifying common indicators across both studies.

## 2. Method

### *2.1 Participants*

All participants reported normal or corrected-to-normal vision and audition and should have at least 3 years of driving experience, be right-handed, and be aged from 21 to 45 years old. Concerning the study 1, twenty participants were involved (11 males, age = 27.15, SD=4.65). The second experiment involved 71 participants (mean age = 28.263 years, SD = 10.585, 39 women). The research protocol was approved by the ethics committee of the Université Gustave Eiffel.

### *2.2 Driving scenario*

In study 1, participants started the experiment with a 5-minute resting time (baseline). Then they drove on a driving simulator on both a highway and rural roads. Initially, they drove manually for two minutes before transitioning to an automated driving system. During the automated phase, they engaged in various tasks: listening to audio that relayed impartial information such as narratives about nature or technological advancements (listening neutral condition) 2) or hearing accounts involving grief, such as stories of individuals losing a loved one to suicide (listening sadness condition), 3) or devising a story from three randomly selected words out of twenty options, each word printed on a card (e.g., 'piano', 'castle', 'robot', 'school') (Wandering neutral condition), 4) or had to recall the saddest episode they ever lived on a paper sheet (Wandering sadness condition).

In study 2, participants were divided into two groups: neutral and sadness group. After a 5-minute resting time (baseline), participants had to listen to podcasts depicting sad stories (sadness induction) or audio about nature and technological advances (neutral condition). Then they had to drive on the simulator on a highway in autonomous mode for 8 minutes.

### *2.3 Measures and Analysis*

Between each step (baseline, induction, and driving) they had to complete some questionnaires to assess their mood states. This procedure allowed to labeling of the physiological dataset according to the participants' mood state.

Physiological data for both studies were collected using the Biopac MP160 system. In study 1, cardiac activity (ECG) and respiratory activity were assessed.

For both studies, we normalized all physiological data by subtracting experimental data from baseline ones. This method adheres to the established norm for examining physiological responses induced by tasks, as evidenced in prior research (e.g., Mazeres et al., 2019).

For study 1, data were analyzed using the last 1 minute of the driving session. For study 2, experimental data correspond to the first 4 minutes of the mood induction procedure. For both studies, a 30-second time window was defined for all physiological data.

For both studies, to classify the drivers' emotional states (neutral or sad), we constructed two classification and regression tree (CRT) models using physiological and psychological indicators.

## 3. Results

### *3.1 Study 1*

We had two mood induction procedures (MIP). One used emotionally connoted podcasts (listening task) and one used an imaginary task (wandering).

The classification model tried to predict sadness according to the MIP. The selected indicators were composed of Heart rate, Heart rate variability (RMSSD and SDNN), Maximum breathing rate, Minimum breathing rate, Mean breathing rate, and Mean breathing amplitude.

| | CRT | Wandering | Listening | % correct |
|---|---|---|---|---|
| **Learning** | **Wandering** | 350 | 59 | 85.6% |
| | **Listening** | 57 | 343 | 85.8% |
| | **Reliability** | 50.3% | 49.7% | **85.7%** |
| **Test** | **Wandering** | 57 | 22 | 72.2% |
| | **Listening** | 29 | 75 | 72.1% |
| | **Reliability** | 47% | 53% | **72.1%** |

Table 1: CRT classification of sadness according to induction method

The results (Table 1) indicate an 85.7% correct classification rate during the learning phase and a 72.1% correct classification rate during the test phase.

| | CRT | Neutral | Sadness | % correct |
|---|---|---|---|---|
| **Learning** (80% data set) | **Neutral** | 190 | 25 | 88.4% |
| | **Sadness** | 40 | 161 | 80.1% |
| | **Reliability** | 55.3% | 44.7% | **84.4%** |
| **Test** (20% data set) | **Neutral** | 27 | 9 | 75.0% |
| | **Sadness** | 11 | 43 | 79.6% |
| | **Reliability** | 42.2% | 57.8% | **77.8%** |

Table 2: CRT decision tree results for sadness detection

Another CRT model tried to correctly predict the emergence of sadness. The results (Table 2) indicated a correct classification of 84.4% in the learning phase and 77.8% in the test phase. The classes created therefore enabled a good identification of a sad state using physiological measures alone.

These results reveal the possibility of using different physiological cues to determine both the driver's activity (listening or not) and emotional state (sad or neutral).

### *3.2 Study 2*

The objective was to predict the drivers' mood state (Neutral or Sad). The chosen indicators included heart rate, SDNN, max heart rate, mean breathing rate, respiratory variability, mean inhale and exhale times, and mean breath amplitude.

| | CRT | Neutral | Sadness | % correct |
|---|---|---|---|---|
| **Learning** | **Neutral** | 153 | 22 | 70.6% |
| | **Sadness** | 15 | 162 | 91.5% |
| | **Reliability** | 37.5% | 62.7% | **81.8%** |
| **Test** | **Neutral** | 32 | 23 | 58.2% |
| | **Sadness** | 7 | 24 | 77.4% |
| | **Reliability** | 47% | 53% | **65.1%** |

Table 3: Number of correct and incorrect detections of sadness based on the actual and classified classes

The model achieved a successful identification rate of 91.5% for truly positive samples (indicating sadness) during the learning phase and 77.4% during the test phase.

Additionally, the model exhibited 81.8% reliability during the learning phase but only 65.1% during the test phase. We used 80% of the data set for the learning phase and 20% for testing. Overall, these results affirm the effectiveness of our model in classifying sadness.

## 4. Discussion

These studies demonstrated the potential for classifying sadness using physiological data during automation. Further research with additional data is needed to enhance classification accuracy. Monitoring the attentional state and emotions of drivers is crucial for road safety, a fact that is increasingly recognized by influential bodies such as the European New Car Assessment Programme (Euro NCAP). Euro NCAP's protocols are evolving to prioritize technologies that can detect driver inattention, drowsiness, and emotional state, which are significant factors in road accidents. For instance, the integration of Occupant Status Monitoring (OSM) technologies into Euro NCAP's Safety Assist protocols is a testament to the importance placed on driver monitoring systems (DMS). These systems are designed to alert drivers when inattention or emotional distress is detected and, if necessary, take control to prevent accidents. The roadmap for future OSM protocol development by Euro NCAP suggests a comprehensive approach to addressing a range of safety risks, including cognitive distraction and the requirements for driver engagement with assisted and automated driving features. This shift not only enhances vehicle safety ratings but also encourages manufacturers to incorporate advanced DMS in their vehicles, ultimately leading to safer roads for everyone. The commitment of Euro NCAP to these advancements underscores the critical role of attentive driving and the profound impact that monitoring systems can have on reducing road fatalities and injuries.

## 5. Conclusions

These results contribute to the understanding of the possibility of monitoring mental states such as sadness. The current systems are designed for laboratory use, and further technological advancements are necessary to facilitate the acquisition of cardiac and respiratory data in a manner suitable for integration into motor vehicles.

## 6. Acknowledgments

These studies were carried out as part of the national projects Autoconduct (ANR-16-CE22-0007) and Autoconduct (AMI AURA 2019). I extend my gratitude to all the researchers who contributed to these endeavors, including A. Fort, H. Tattegrain, A. Hidalgo-Muñoz, A. Béquet, and the engineers D. Ndiaye, F. Moreau, B. Richard.

## 7. References

Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience*, *17*(11), 718-731. https://doi.org/10.1038/nrn.2016.113

Garrity, R. D., & Demick, J. (2001). Relations among personality traits, moods states, and driving behaviors. *Journal of Adult Development*, *8*(2), 109-118.

Gruberger, M., Ben-Simon, E., Levkovitz, Y., Zangen, A., & Hendler, T. (2011). Towards a Neuroscience of Mind-Wandering. *Frontiers in Human Neuroscience*, *5*, 56. PMC. https://doi.org/10.3389/fnhum.2011.00056

Jallais, C., Gabaude, C., & Paire-Ficout, L. (2014). When emotions disturb the localization of road elements: Effects of anger and sadness. *Transportation Research Part F: Traffic Psychology and Behaviour*, *23*(0), 125-132. https://doi.org/10.1016/j.trf.2013.12.023

Kam, J. W. Y., Rahnuma, T., Park, Y. E., & Hart, C. M. (2022). Electrophysiological markers of mind wandering: A systematic review. *NeuroImage*, *258*, 119372. https://doi.org/10.1016/j.neuroimage.2022.119372

Lagarde, E., Chastang, J.-F., Gueguen, A., Coeuret-Pellicier, M., Chiron, M., & Lafont, S. (2004). Emotional Stress and Traffic Accidents: The impact of separation and divorce. *Epidemiology*, *15*(6), 762-766.

Mesken, J., Hagenzieker, M. P., Rothengatter, T., & de Waard, D. (2007). Frequency, determinants, and consequences of different drivers' emotions: An on-the-road study using self-reports, (observed) behaviour, and physiology. *Transportation Research Part F: Traffic Psychology and Behaviour*, *10*(6), 458-475. https://doi.org/10.1016/j.trf.2007.05.001

Nilsson, E. J., Bärgman, J., Ljung Aust, M., Matthews, G., & Svanberg, B. (2022). Let Complexity Bring Clarity: A Multidimensional Assessment of Cognitive Load Using Physiological Measures. *Frontiers in Neuroergonomics*, *3*, 787295. https://doi.org/10.3389/fnrgo.2022.787295

Smallwood, J., & Schooler, J. W. (2015). The science of Mind Wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology*, *66*, 487-518.

Stemmler, G., Aue, T., & Wacker, J. (2007). Anger and fear: Separable effects of emotion and motivational direction on somatovisceral responses. *International Journal of Psychophysiology*, *66*(2), 141-153. https://doi.org/10.1016/j.ijpsycho.2007.03.019

Stephens, A. N., & Groeger, J. A. (2011). Anger-congruent behaviour transfers across driving situations. *Cognition & Emotion*, *25*(8), 1423-1438. https://doi.org/10.1080/02699931.2010.551184

# Detecting physiological and physical markers of driver passive fatigue for autonomous driving

Stephanie Dabic [1*], Antoine Favre-Felix[2], Laora Kerautret [1]

[1] Valeo Switches & Smart Controls Product Line , Rue Jules VERNE - Vetraz Monthoux
BP 509 - 74106 Annemasse Cedex, FRANCE
[2] Innovateam – 52, Quai de Dion Bouton 92800 PUTEAUX
[*].stephanie.dabic@valeo.com

**Abstract: Passive fatigue (PF) induced during autonomous vehicle operation poses a significant safety concern, particularly during transitions to take over control. In this study, we conducted a randomized controlled experiment utilizing a high-fidelity driving simulator with 68 participants. Employing a between-group design, PF was assessed using the Karolinska Sleepiness Scale (KSS), ocular and cardiac parameters. Our aim was to find the physiological and physical signature of PF during autonomous driving. Preliminary findings on fatigue are presented to identify physiological markers of PF preceding behavioral manifestations on pupil size, Vertical Gaze Variance, Very Low Frequency (VLF) and Low Frequency (LF) power parameters. These results will be turned into recommendations for the development of an embedded system to mitigate PF during partially autonomous driving.**

## 1. Introduction

Offering guidelines for creating an embedded system aimed at reducing passive fatigue (PF) during partially autonomous driving has become a major issue for safety. Indeed, fatigue and the subsequent decrease in vigilance are recognized as key factors contributing to traffic accidents and bad consequences (Nason, 2005; Thiffault, Bergeron, 2003). Several studies highlight that both excessive and insufficient cognitive demands can impair performance. Hancock and Desmond (2001) proposed a cognitive fatigue model distinguishing between Active Fatigue (AF) from excessive demands and Passive Fatigue (PF) from insufficient stimulation. While reducing workload may ease AF, PF requires different strategies to be reduced. In driving, AF stems from continual task adjustments, while PF arises from minimal perceptual-motor demands (Bernhardt et al., 2019; Saxby et al., 2013). Extended periods of autonomous driving can induce PF, posing safety risks during manual control handover.

Moreover, vigilance decrement is the most robust effect of fatigue and sleepiness (Dingues, 1995) .Different definitions of concepts like vigilance, fatigue, arousal, and activation may cause confusion. It's important to clarify their meanings. Vigilance can be seen in two ways: physiological processes of alertness or wakefulness, and information processing with sustained attention (Körber and als, 2015) . Fatigue is a broad term covering both physiological and psychological aspects.

Nevertheless heart rate (HR) and heart-rate variability (HRV) are generally considered to be good relative indicators of workload/fatigue, but effects can be inconsistent. Indeed, studies show that HR increases and HRV decreases during demanding mental processing, although it has also been shown that HR decreases significantly during a monotonous driving task (Larue et als, 2011). Lee et al. (1990) pointed out that changes in HR mainly reflect physical fatigue, while HRV can comprehensively reflect physical fatigue and mental fatigue. Dou (2017) found that the HRV index of Very Low Frequency (VLF) is suitable for predicting driving fatigue. index.

Körber et al. (2015) found that PF was induced when participants were monitoring a driving simulator with automation capabilities that controlled longitudinal and lateral steering, as indexed by a continual overall decrease in pupil diameter, and a general trend of slower reaction times on an auditory oddball task.

Jin and Yu (2018) found that pupil diameter could be an effective measure of operator fatigue: under PF, pupil diameter decreased significantly while under active fatigue, pupil diameter increased. The two forms of fatigue produce antagonism and jointly restrict the change of pupil diameter. Therefore, the degree of fatigue cannot be determined simply by the reduction in pupil diameter. One other indicator seen to be interesting to investigate : Percent Road Center (PRC) defined as the proportion of time that a driver's eyes are focussed on the road center (Victor, 2005).

The aim of this study is to find pertinent physiological and physical parameters to detect PF induced in a simulator context during automated system supervision.

*Setup*

A total of 68 participants were enrolled in this study. All participants held a valid driver's license for at least one year and were employees of Valeo. Random assignment placed participants into one of five experimental groups, which will be detailed later in the paper.

The experiment took place at the Valeo Annemasse site, utilizing a driving simulator (Figure 1). The simulator features two car seats, an automatic gearbox, and Logitech G2 steering wheel and pedals.

*Figure 1. Illustration of simulator setup*

Unity 3D software was used to design the homemade simulated driving environment.

ECG was recorded with a BIOPAC MP160 system at 500 Hz and the eye-tracking data was recorded with a FOVIO system at 62 Hz. The HRV parameters and the eye-tracking parameters were calculated over 5-minute windows.

### 1.1 *Driving situation and procedure :*

All drives were highway-routes (with 3 lanes in each direction) with very little traffic at 110 km/h on both sides of the highway on Unity Homemade software. This road and traffic display was chosen in order to create a monotonous task with low demands. After 40 minutes of autonomous driving, a take-over request was initiated by the system through a distinct auditory and visual warning signal "take the hand". Participants then drove 5 minutes in a manual mode after successfully handling the safety hazard.

Driver drowsiness was subjectively measured using the Karolinska Sleepiness Scale (KSS), a 9-point scale (Åkerstedt & Gillberg, 1990). KSS was administered before and after the driving scenario and orally every 5 minutes during the driving scenario.

In order to avoid confounding effects (such as those described in the May and Bradley model described above), we made a sleep diary of the participants over the two days before the test in order to avoid retaining participants with sleep deprivation.

## 2. Results

Statistical analyses were conducted to find the signature of PF during autonomous driving.

### 2.1 *KSS (subjective data)*

We performed a Linear Mixed Model (LMM) for each dependent variable, with a 5-minute time segment as fixed effects. In order to account for a general variability across drivers, participant ID was included in random effects. Time factor was a within subject factor (every 5 minutes the KSS, and one before and after tests).

We observed the main effect of Time *F( 23.369, p <.001)*. In other words, the level of subjective PF increased variously over time. Post-Hoc analysis (with Bonferroni correction) showed significant differences from 15 to 30 min compared to first segment (Figure 2).

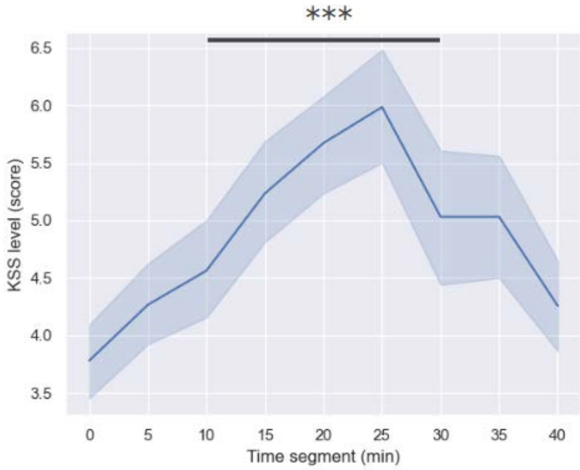



***Figure 2 : Illustration of the timeline's procedure set up for KSS data***

### 2.2 *Eye-tracking data*

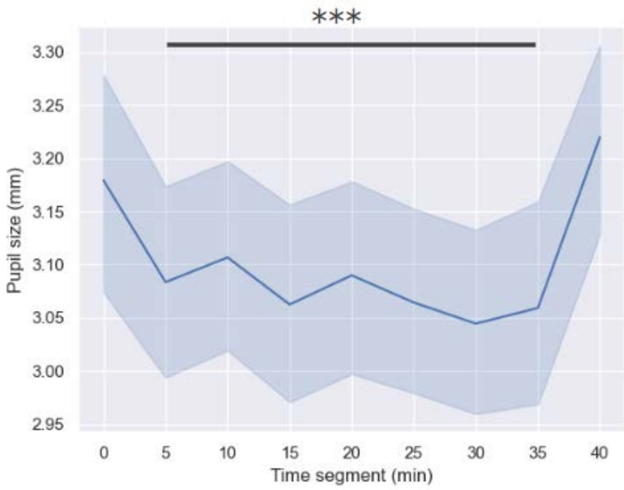



***Figure 3 : Illustration of the timeline's procedure set up for Pupil size***

We performed a LMM on eye-tracking data. We observed the main effect of Time *F( 19.075, p <.001)*. In other words, pupil size decreased variously over time . Post-Hoc analysis (with Bonferroni correction) showed significant differences from 5 to 35 min compared to first segment (Figure 3).

We observed the main effect of Time F( 14.922, p <.001). In other words, Vertical Gaze varied less over time than at the beginning . Post-Hoc analysis (with Bonferroni correction) showed significant differences from 10 to 40 min compared to first segment (Figure 4).

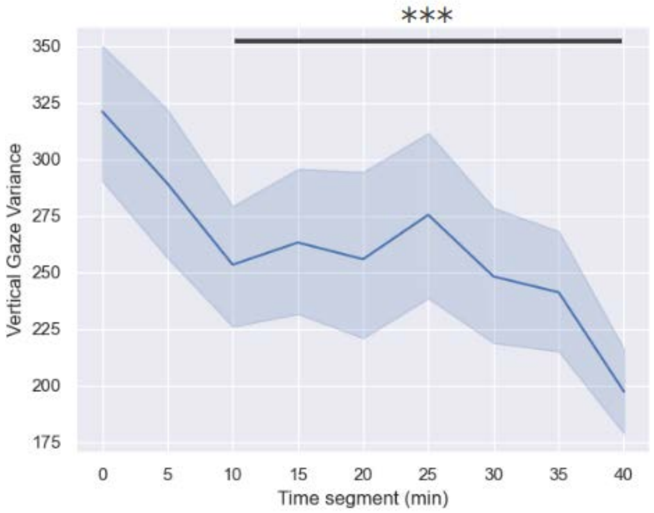



***Figure 4 : Illustration of the timeline's procedure set up for Vertical Gaze Variance in cm²***

### 2.3 *Cardiac data*

We performed a LMM on HRV data. We observed the main effect of Time *F( 8,472, p <.001)*. In other words, VLF increased over time. Post-Hoc analysis (with Bonferroni correction) showed significant differences from 25 to 35 min compared to first segment (Figure 5).

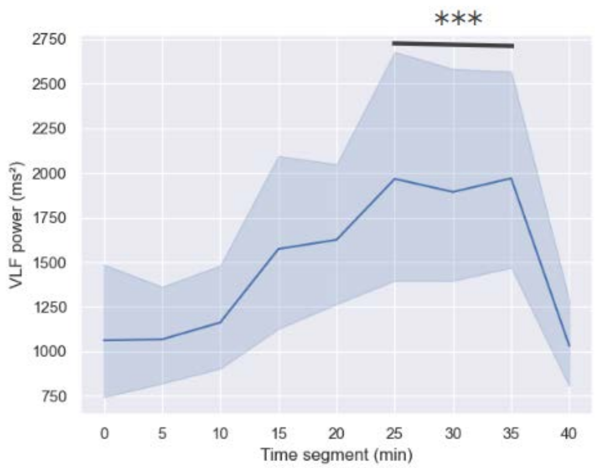



***Figure 5 : Illustration of the timeline's procedure set up for VLF power***

We performed a LMM on HRV data. We observed the main effect of Time *F( 8,472, p <.001)*. In other words, LF increased over time. Post-Hoc analysis (with Bonferroni correction) showed significant differences from 15 to 35 min compared to first segment (Figure 6).

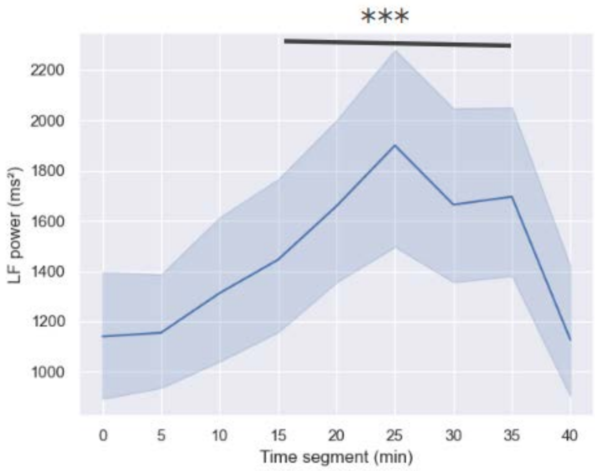



***Figure 6 : Illustration of the timeline's procedure set up for LF power***

## 3. Discussion

In this study we explored some physiological and physical parameters of PF.

First, KSS has been widely used as a self-reported measure of drowsiness and PF in studies of shift work, jet lag, attention and performance, and clinical settings (for a review, see Kaida et al., 2006). We can say PF was induced considering the increasing score of the KSS. Participants felt more tired from 10 to 30 min significantly without real sleep deprivation.

Moreover, we found that pupil size decreased significantly with PF. This result is coherent with Körber and als (2015) study showing significant effects on eye-tracking parameters, especially decrease on pupil diameter corresponding to vigilance decrement and increasing mind wandering. In effect authors showed this fatigue signature occurs without active task engagement.

Nevertheless, the variation in pupil diameter alone can be an indicator of many other things, which is why it is important to complement it with other visual parameters such as VGV, which shows the driver's involvement in the driving task and the supervision of autonomous driving. This visual indicator is also relevant as a marker of the onset of PF and of decrease of vigilance of drivers (Kim and all, 2023).

Finally, PF could be shown with increasing VLF and LF power parameters as previous results (Henelius et al., 2014) , and have been associated with decreased vigilance caused by total or partial sleep deprivation (Chua et al., 2012).

## 4. Conclusions and perspectives

These findings will be translated into suggestions for developing an embedded system aimed at detecting and reducing PF during partially autonomous driving with physical and physiological data and training algorithms for individual detection. All of them need to be combined to have the best detection method.

## 5. Acknowledgments

The analysis of this research was conducted with French financial support of BPI in the SERENITE project.

## References


Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. International journal of neuroscience, 52(1-2), 29-37.

Bernhardt, K. A., Kelley, A. M., Feltman, K. A., & Curry, I. P. (2019). Rest and activity patterns of Army aviators in routine and operational training environments. Aerospace Medicine and Human Performance, 90(1), 48-52.

Chua, E. C. P., Tan, W. Q., Yeo, S. C., Lau, P., Lee, I., Mien, I. H., ... & Gooley, J. J. (2012). Heart rate variability can be used to estimate sleepiness-related decrements in psychomotor vigilance during total sleep deprivation. Sleep, 35(3), 325-334.

Dinges, D.F., 1995. An overview of sleepiness and accidents. J. Sleep Res. 4, 4–14.

Dou Guangbo. (2017). Behavioral Analysis and Measurement of Passive Fatigue of Drivers. Doctoral Dissertation, Dalian: Liaoning Normal University.

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

Henelius, A., Sallinen, M., Huotilainen, M., Müller, K., Virkkala, J., & Puolamäki, K. (2014). Heart rate variability for evaluating vigilant attention in partial chronic sleep restriction. Sleep, 37(7), 1257-1267.

Jin, H., Yu, G., & Liu, H. (2018). Effectiveness analysis of pupil diameter detection for air traffic controller's fatigue. Journal of Beijing University of Aeronautics and Astronautics, 44(7), 1402.

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. Clinical neurophysiology, 117(7), 1574-1581.

Kim, D., Park, H., Kim, T., Kim, W., & Paik, J. (2023). Real-time driver monitoring system with facial landmark-based eye closure detection and head pose recognition. Scientific reports, 13(1), 18264.

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.

Lee, B. W., Ferreira, J. M., Ghamaty, S., Yang, K. N., & Maple, M. B. (1990). Magnetic Ordering of Rare Earth (R) Ions in RBa 2 Cu 3 O 7− δ Compounds (R= Nd, Sm, Dy, Er; 0≤ δ≤ 1). In Oxygen Disorder Effects in High-Tc Superconductors (pp. 151-162). Springer, Boston, MA.

Nason, N. R. (2005). TRAFFIC SAFETY FACTS 2005-A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System, National Highway Traffic Safety Administration. National Center for Statistics and Analysis, US Department of Transportation, Washington, DC, 20590. TRAFFIC SAFETY FACTS 2005-A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System, National Highway Traffic Safety Administration. National Center for Statistics and Analysis, US Department of Transportation, Washington, DC, 20590.

Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. Accident Analysis & Prevention, 35(3), 381-391.

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.

Victor, T. (2005). Keeping eye and mind on the road (Doctoral dissertation, Acta Universitatis Upsaliensis).

# Still fit to drive? – How Car Sickness affects Takeover and Driving Performance

Myriam Metzulat [1*], Barbara Metz [1], Wilfried Kunde [2]

[1] Würzburger Institut für Verkehrswissenschaften (WIVW) GmbH, Robert-Bosch-Straße 4, 97209 Veitshöchheim, Germany
[2] Department of Psychology (III), University of Würzburg, Röntgenring 11, 97070 Würzburg, Germany
[*]metzulat@wivw.de

**Abstract: In automated driving the driver is likely to experience car sickness when not paying attention to the road. Considering previous findings of cognitive performance decrements due to motion sickness, and the fact that driving demands cognitive resources, the aim of this study was to understand if and how car sickness affects takeover and subsequent driving performance. *N* = 33 participants took part in this test-track study with a wizard-of-Oz vehicle including a car sickness condition and a control condition without experiencing car sickness in randomised order. In both conditions four takeover requests were triggered, after which participants performed four different driving tasks. Subjective measures of car sickness, criticality, mental workload, and performance were recorded, as well as objective performance and driving measures. Results of the subjective data showed that takeovers and driving with car sickness were perceived as significantly more critical and demanding compared to without, and that most participants felt impaired by car sickness while driving. These results will be validated with the objective data.**

## 1. Introduction

With the introduction of automated driving (SAE Level 3 and above; SAE, 2021), the role of the driver will change to that of a (part-time) passenger with the possibility of engaging in non-driving related tasks (NDRTs). Among passengers, especially when performing NDRTs, car sickness, a subtype of motion sickness, is a common phenomenon (Rolnick & Lubow, 1991; Schmidt et al., 2020). Therefore, the risk of experiencing car sickness in automated driving will increase, resulting in a potentially car sick driver when taking over control of the vehicle.

Previous studies on the effect of motion sickness on task performance have revealed impairments in cognitive performance, such as increased reaction times (Bos, 2015; Smyth et al., 2019), impaired hand/arm coordination (Smyth et al., 2019), reduced performance on visual search (Golding & Kerguelen, 1992) and perception task (Kaplan et al., 2017). Regarding the effect on driving performance, studies on the effects of simulator sickness have shown prolonged braking reaction times (Reinhard, Tutulmaz, et al., 2019) and reduced average speed (Gálvez-García et al., 2020; Reinhard, Kleer, et al., 2019). However, to the authors' knowledge there are no studies investigating the effects of car sickness on driving performance in the real world in context of automated driving including takeover performance. The previously reported performance degradations under motion sickness could lead to safety critical behaviour in complex takeover situations, such as construction sites or obstacles ahead, which require a quick understanding of the situation, as well as appropriate decisions and reactions. Therefore, the aim of this study was to investigate the effects of car sickness on takeover and driving performance in a real-world setting.

## 2. Method

### *2.1 Experimental design*

A wizard-of-Oz test track study was conducted in which participants experienced two test conditions on separate days and in a randomised order. Symptoms of car sickness were induced in the car sickness condition but not in the control condition. In both conditions, participants had to take over vehicle control four times and perform the four following driving tasks in randomised order after each takeover while sitting on the driver's seat: slalom with a fixed speed (25 km/h), slalom with a freely chosen speed, target braking and two emergency stops upon a warning tone. The target braking required the driver to accelerate to 30km/h and then to stop as accurately as possible at the target position.

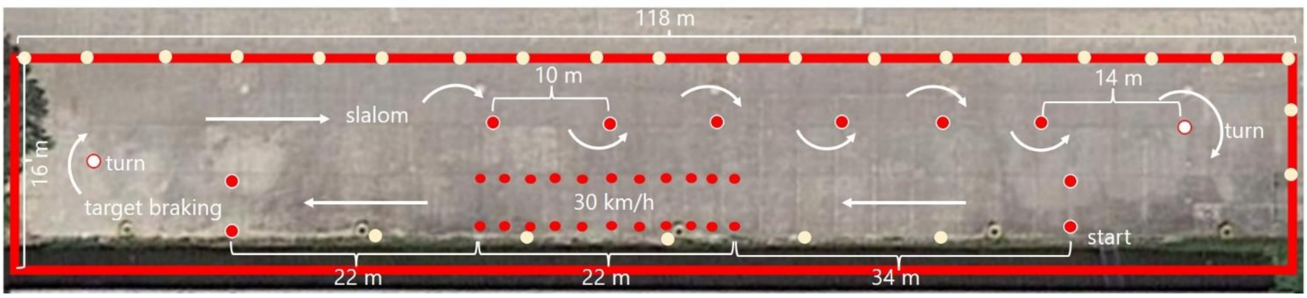



***Fig.1.*** *Test track with pylons set up to mark the driving tasks (emergency braking on a random spot on each straight)*

*2.2 Test environment and set-up*

The study took place on a closed test track (see Fig. 1). The wizard vehicle was an AUDI Q7 with a conventional steering wheel and pedals on the front left. In addition, the vehicle had instructor pedals on the front right and a joystick to steer the vehicle. The wizard could trigger a takeover request (TOR). Participants had to press a button to confirm the takeover before they could drive themselves.

*2.3 Test conditions and procedure*

To induce car sickness in the car sickness condition, the wizard simulated a dynamic automated ride including a stop-and-go scenario on the one straight of the test track and a slalom back on the other straight. Participants had to play a maze game (Cabbiegames, 2022) on the smartphone during the ride to increase the likelihood of car sickness (see Fig. 2).

***Fig.2.** Participant engaging in the NDRT and wizard driving with the joystick on armrest (arrow)*

Every 30 s, participants had to verbally indicate their current level of car sickness on the Misery Scale (Bos et al., 2005). The TOR was triggered after 10 minutes or when the participant indicated a MISC level of 7 (mild to moderate nausea). In the control condition, the participants did not experience an automated drive but played the game in standstill for the same time and the wizard drove off a few meters to trigger the TOR. After taking over, participants had to perform one of the driving tasks. Afterwards, a break of at least 6 minutes began, in which participants completed the MSAQ (Gianaros et al., 2001) and questionnaires regarding the takeover and driving task. After the break, the same procedure was repeated for the remaining three driving tasks.

*2.4 Test sample*

33 subjects (17 female), who were selected based on their moderate to severe susceptibility to car sickness, participated in the study. The mean age was 41.9 years (*SD* = 15.5).

*2.5 Dependent measures*

To measure the subjective criticality of each takeover and driving task, the Criticality Scale by Neukum and Krüger (2003) from 1=harmless to 10=uncontrollable was used. To assess the mental workload of each driving task an adapted version of the NASA-TLX Questionnaire was applied (Hart & Staveland, 1988).

In a short interview at the end, the participants were asked whether they felt impaired when driving and whether they had adapted their driving behaviour due to car sickness. In addition, objective measures of the takeover and driving performance as well as driving dynamics were recorded, which were still under analysis at the time the abstract was submitted.

## 3. Results

In 78.5% of all car sickness rides, a MISC level of 7 was reached. Accordingly, the MSAQ scores were significantly higher for all driving tasks in the car sickness condition compared to the control condition (manipulation check). All statistic values of the conducted t-tests can be found in Table 1. The criticality of all four takeovers and of three driving tasks (all except for slalom free) differed significantly between conditions with higher ratings in the car sickness condition (see Fig. 3). In both conditions, the average criticality of takeovers was assessed within the range of harmless.

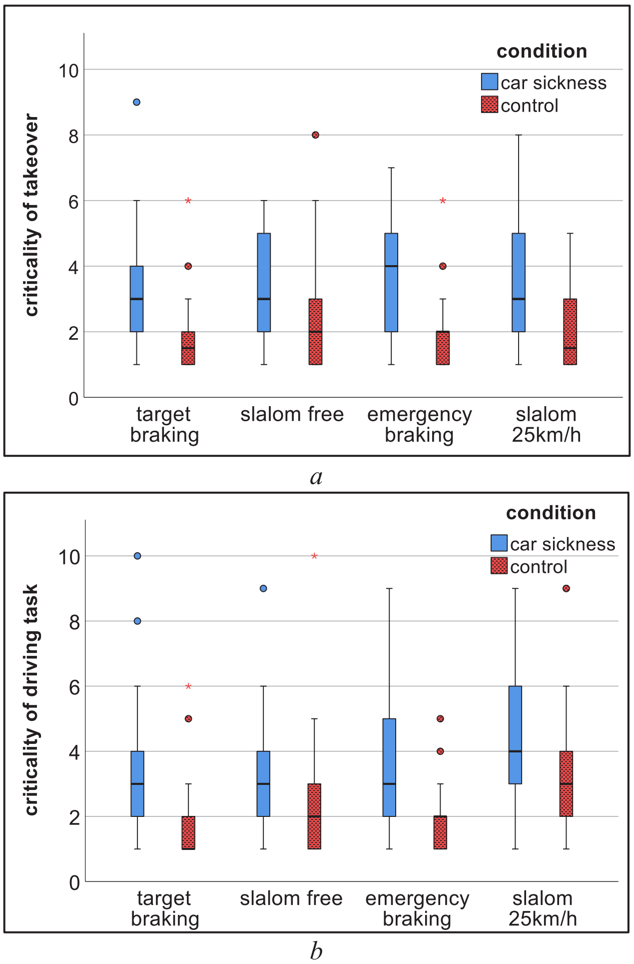



***Fig. 3.** Criticality per driving task for both conditions **(a)** Of takeovers, **(b)** Of driving tasks*

Mental workload was rated significantly higher for all driving tasks with car sickness than without. 72.7% of all participants reported feeling impaired by symptoms of car sickness when performing the driving tasks. Cognitive impairments were mentioned most frequently, such as decreased concentration, orientation and attention or slower

***Table 1:*** *Statistics of t-tests and descriptives of MSAQ, NASA-TLX and criticality of takeover and driving task*

| Task | Measure | t | p | *M (SD)* car sickness | *M (SD)* control |
|---|---|---|---|---|---|
| target braking | MSAQ* | $t(32) = 9.94$ | <.001 | 40.27 (15.39) | 13.71 (3.52) |
| | NASA-TLX | $t(32) = 5.60$ | <.001 | 33.56 (16.84) | 20.93 (12.27) |
| | criticality takeover | $t(32) = 4.48$ | <.001 | 3.30 (1.83) | 1.91 (1.31) |
| | criticality driving task | $t(32) = 3.57$ | .001 | 3.18 (2.19) | 1.70 (1.19) |
| slalom free | MSAQ* | $t(31) = 10.91$ | <.001 | 40.63 (14.28) | 13.99 (3.71) |
| | NASA-TLX | $t(31) = 3.06$ | .005 | 33.28 (17.39) | 26.39 (14.34) |
| | criticality takeover | $t(31) = 3.74$ | .001 | 3.28 (1.55) | 2.22 (1.68) |
| | criticality driving task | $t(31) = 1.73$ | .094 | 3.16 (1.89) | 2.50 (1.70) |
| emergency braking | MSAQ* | $t(31) = 9.62$ | <.001 | 39.92 (15.20) | 14.66 (5.05) |
| | NASA-TLX | $t(31) = 4.36$ | <.001 | 37.54 (21.83) | 22.57 (8.42) |
| | criticality takeover | $t(31) = 4.99$ | <.001 | 3.67 (1.90) | 2.00 (1.32) |
| | criticality driving task | $t(31) = 5.14$ | <.001 | 3.73 (2.27) | 1.78 (0.98) |
| slalom 25km/h | MSAQ* | $t(32) = 9.95$ | <.001 | 41.15 (15.27) | 14.44 (5.16) |
| | NASA-TLX | $t(32) = 4.45$ | <.001 | 40.91 (18.90) | 28.73 (17.65) |
| | criticality takeover | $t(32) = 4.40$ | <.001 | 3.48 (1.87) | 1.94 (1.14) |
| | criticality driving task | $t(32) = 3.41$ | .002 | 4.27 (2.28) | 3.03 (1.96) |

**possible min = 11.11*

reaction time. 69.7% of all participants stated that they have adapted their driving behaviour due to car sickness. A lower speed and a more defensive driving style were the most often stated adaptions.

## 4. Discussion

The results on the level of car sickness showed that the manipulation has worked, i.e. car sickness was considerably induced. All takeovers and driving tasks with speed or braking instructions were rated significantly more critical when driving under car sickness. In the slalom, where participants were free to choose their speed, criticality was not significantly higher with carsickness. It is likely that participants adjusted their speed to make this driving task less demanding. Some participants consistently stated that they reduced their speed and drove more defensively due to car sickness, which is in line with the results of the simulator studies mentioned above (see 1. Introduction). Safety is unlikely to be compromised by reducing speed and driving defensively in an appropriate manner. However, a significant proportion of the sample felt impaired by car sickness while driving, mainly due to reduced cognitive performance, like e.g., slower reaction time or reduced attention. These findings are consistent with previous findings on the effects of motion sickness on performance (e.g., Bos, 2015; Smyth et al., 2019). Such impairments can in fact be safety critical if, for example, obstacles are detected too late or the braking is delayed in emergency situations. However, so far, the results of the present study are based on subjective reports only. It is of interest to check whether the participants' subjective assessments are reflected in the objective measurements. Results on this aspect will be available at the time of the conference and can be presented.

## 5. Conclusions

In conclusion, driving with symptoms of car sickness is perceived as more demanding, critical and impairs subjective driving performance. Further analysis will be done to verify these findings with objective performance measures.

## 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. Neither the European Commission nor CINEA – in its capacity of Granting Authority – can be made responsible for any use that may be made of the information this document contains. The authors would like to thank all partners within Hi-Drive for their cooperation and valuable contribution.

## References

Bos, J. E. (2015). Less sickness with more motion and/or mental distraction. *J Vestib Res*, *25*(1), 23-33.

Bos, J. E., MacKinnon, S. N., & Patterson, A. (2005). Motion sickness symptoms in a ship motion simulator: effects of inside, outside, and no view. *Aviat Space Environ Med*, *76*(12), 1111-1118.

Cabbiegames. (2022). *Classic Labyrinth 3d Maze*. Retrieved 09.02.2023 from http://www.classiclabyrinth.com/index.html

Gálvez-García, G., Aldunate, N., Bascour–Sandoval, C., Barramuño, M., Fonseca, F., & Gómez-Milán, E. (2020). Decreasing motion sickness by mixing different techniques. *Appl Ergon*, *82*, 102931.

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. *Aviat Space Environ Med*, *72*(2), 115 - 119.
Golding, J. F., & Kerguelen, M. (1992). A comparison of the nauseogenic potential of low-frequency vertical versus horizontal linear oscillation. *Aviat Space Environ Med*, *63*(6), 491-497.
Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139-183). Elsevier.
Kaplan, J., Ventura, J., Bakshi, A., Pierobon, A., Lackner, J. R., & DiZio, P. (2017). The influence of sleep deprivation and oscillating motion on sleepiness, motion sickness, and cognitive and motor performance. *Autonomic Neuroscience*, *202*, 86-96.
Neukum, A., & Kruger, H.-P. (2003). Fahrerreaktionen bei Lenksystemstörungen–Untersuchungsmethodik und Bewertungskriterien. *VDI BERICHTE*, *1791*, 297-318.
Reinhard, R. T., Kleer, M., & Dreßler, K. (2019). The impact of individual simulator experiences on usability and driving behavior in a moving base driving simulator. *Transportation research part F: traffic psychology and behaviour*, *61*, 131-140.
Reinhard, R. T., Tutulmaz, E., Rutrecht, H. M., Hengstenberg, P., Geissler, B., Hecht, H., & Muttray, A. (2019). Effects of visually induced motion sickness on emergency braking reaction times in a driving simulator. *Hum Factors*, *61*(6), 1004-1018.
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.
SAE. (2021). J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. *SAE International/ISO*. www.sae.org
Schmidt, E. A., Kuiper, O. X., Wolter, S., Diels, C., & Bos, J. E. (2020). An international survey on the incidence and modulating factors of carsickness. *Transportation research part F: traffic psychology and behaviour*, *71*, 76-87.
Smyth, J., Birrell, S., Mouzakitis, A., & Jennings, P. (2019). Motion sickness and human performance–exploring the impact of driving simulator user trials. Advances in Human Aspects of Transportation: Proceedings of the AHFE 2018 International Conference on Human Factors in Transportation, 2018, Orlando, Florida, USA.

# Pre-crash driver behaviors in motor vehicle crashes caused by sudden illness

Emma Nilsson [1*], Mirta Zelenika Zeba [1], Emma Tivesten [1,2]

[1] Volvo Cars Safety Centre, Volvo Cars, Gothenburg, Sweden
[2] Chalmers University of Technology, Gothenburg, Sweden
*emma.j.nilsson@volvocars.com

**Abstract: Sudden illness can cause the most extreme form of disconnected drivers by rapidly and unexpectedly impairing drivers' ability to drive. Consequently, it poses a serious safety concern in traffic. However, studies in this field are few and knowledge limited. Therefore, this study will investigate drivers' pre-crash vehicle handling and signs of illness in car crashes caused by sudden illness. Crashes will be identified in Volvo Cars' Accident Database through a word search. Case files will be derived from pre-crash vehicle signal recordings, in-depth crash investigation reports, and mail survey questionnaires. Descriptive statistics (on, e.g., driving contexts and crash types) will be conducted for all identified cases. Furthermore, an explorative mixed method analysis will be conducted to identify different types of pre-crash vehicle handling and signs of illness. The findings will inform the development of technologies that can detect sudden illness and prevent accidents.**

## 1. Introduction

Sudden illness, such as epileptic seizures, cardiac arrhythmias, strokes, and hypoglycemic events, can quickly and with little or no forewarning impair drivers' physical as well as cognitive abilities and lead to car crashes.

Research shows that sudden illness is the main contributing factor to a significant proportion of crashes and road fatalities. At least 9-15% of driver fatalities are caused by sudden illness, most commonly ischemic heart disease (Ahlm et al., 2001; Breen et al., 2018; Brodie et al., 2019; Tervo et al., 2008). Fewer studies have investigated how often severe, but not necessarily fatal, crashes are caused by sudden illness. Hanna (2009) found it to be 1.3% of the crashes investigated and estimated that approximately 20 000 crashes annually in the United States are caused by sudden illness. In two similar Australian studies, the proportion of crashes caused by sudden illness was approximately ten times higher: 11.5% (Lindsay & Baldock, 2008) and 14.5% (Fitzharris et al., 2020). Common medical conditions were seizures, cardiac-related events, diabetic reactions, and loss of consciousness for other or unknown reasons.

To prevent crashes caused by sudden illness, knowledge is needed about how to detect medical emergencies in car drivers and how to support these drivers. However, medical emergencies are rare and unpredictable events. Consequently, there are few studies and limited knowledge about typical signs of illness and courses of events in crashes caused by sudden illness.

To date, studies have shown that most crashes caused by sudden illness are single vehicle crashes initiated by lane departures (Brodie et al., 2019; Hanna, 2009; Neal et al., 2018). Furthermore, vehicle pre-crash movements have been described as "out of control" for approximately 50% of seizure-related crashes (Neal et al., 2018).

Case studies reported by Lindsay and Baldock (2008), Marinella (2004), Motozawa et al. (2005), and Sakurai et al. (2014) provide some additional information. Together, these case studies include seven crashes caused by drivers suffering reduced or lost consciousness due to different medical conditions. In one case (Sakurai et al., 2014), there is a video recording of the driver that provides a good understanding of the driver's behavior before the crash. In the remaining cases, the case descriptions provide a good understanding of the medical conditions and crash sites. However, information about the drivers' behaviors (e.g., posture and pedal usage) is limited since there are no vehicle signals, passenger testimonies, or video recordings of the events.

Given the limited knowledge of how sudden illness-related crashes occur, further studies on the subject are needed. The aim of this study is therefore to investigate driver's pre-crash vehicle handling and signs off illness in crashes where sudden illness was the main contributing factor. For this, Volvo Cars' Accident Database will be used. The database contains pre-crash vehicle signals as well as driver and passenger testimonies. This study will thus provide more detailed case descriptions for a larger number of crashes than previously done. This will significantly increase our knowledge in this field.

## 2. Method

This study will provide a retrospective analysis of cases in Volvo Cars' Accident Database. All drivers (except deceased ones) have given their written consent for Volvo Cars to collect and use their data for traffic safety research. Ethical approval has been applied for the study.

### *2.1 Data sources*

Volvo Cars' Accident Database contains over 7000 crash cases involving Volvo cars in Sweden. Most cases include a mail survey questionnaire that the drivers have answered. The questionnaire contains specific questions regarding, for example, the traffic environment and sustained injuries. It also contains two open ended questions with room for descriptions in free text and drawings. The first question asks the driver to explain the events leading up to the crash, and the second question asks for any additional information on why the crash occurred.

In addition, some cases include in-depth investigations that can, for instance, contain on-scene investigations, interviews, vehicle inspections, and medical records.

Some cases also have Event Data Recorder (EDR) data that contains pre-crash vehicle signals from the last five seconds leading up to the crash. Vehicle speed, gas and brake pedal application, and steering wheel angle will be investigated in this study.

### *2.2 Selected data*

Cases where sudden illness was the main contributing factor to the crash will be identified through a word search. The word search will include illness-related terms and be applied to the transcribed free text questionnaire responses and to short case description notes made by the in-depth investigation team. All cases in the database that occurred between 2010-2024 will be included in the word search.

### *2.3 Analysis*

Descriptive statistics will be conducted on driver age and gender, trip duration, road type, posted speed limit, presence of passengers in the car, number and type of road users involved, first critical event, and medical condition. Medical conditions will be classified according to the International Classification of Diseases by medical doctors, based on the available illness descriptions provided by the drivers, passengers, or medical records.

Furthermore, cases with EDR data and/or more detailed descriptions of pre-crash events will be included in an explorative mixed method analysis. The analysis will focus on identifying different types of driver pre-crash vehicle handling (e.g., pedal usage and steering) and signs of illness (e.g., driver posture and self-experienced symptoms), using multiple types of data sources.

## 3. Results

Results will be presented at the conference.

## 4. Discussion

Knowledge about driver behaviors and signs of illness during medical emergencies is needed for the development of safety systems. This study will provide unique information on this topic by studying both pre-crash vehicle signals and testimonies from drivers and passengers. To our knowledge, this is the first time such data has been used on a larger scale to understand crashes caused by sudden illness.

Currently, driver unresponsiveness (which may be caused by sudden illness) is suggested to be detected through eyes off road behavior (see Euro NCAP, 2023). However, other ways of detecting impairment are also needed, and advanced vision and/or biometric sensors are proposed to be used in future systems (Euro NCAP, 2022). Investigating signs of illness in real-life crashes is thus of great importance to inform development of such systems to enable early and accurate detection of driver impairment.

Once impairment has been detected, safety functionality should support drivers to prevent or mitigate crashes. Knowing which behaviors drivers exhibit in crashes caused by sudden illness is then of great importance, as safety system interventions can be affected by the drivers' actions. For example, a distinct steering or pedal input may be interpreted by the system as a deliberate act to override the system. However, in case of sudden illness, this may limit the system's effectiveness since the driver may be unconscious or for other reasons unable to stay in control of driving. An increased understanding of how sudden illness affects driver behavior can therefore improve the effectiveness of driver support systems by tailoring them to the detected driver state.

## 5. Conclusions

To our knowledge, this is the first time that driver vehicle handling and signs of illness will be analyzed for crashes caused by sudden illness, using both pre-crash vehicle signals and driver and passenger testimonies, in a larger dataset. The study findings will hence provide valuable and novel insights into how sudden illness can lead to crashes and how such crashes may be prevented. These insights can inform the development of driver assistance systems for suddenly impaired drivers and further enhance real-life safety.

## References

Ahlm, K., Eriksson, A., Lekander T. and Björnstig, U. (2001). Alla dödsfall i trafiken är inte ”dödsolyckor” – en analys av officiell statistik över dödsfall i svensk vägtrafik år 1999. *Läkartidningen, 17*(98), 2016-2022.

Breen, J. M., Naess, P. A., Gjerde, H., Gaarder, C. and Stray-Pedersen, A. (2018). The significance of preexisting medical conditions, alcohol/drug use and suicidal behaviors for drivers in fatal motor vehicle crashes: a retrospective autopsy study. *Forensic Science, Medicine and Pathology*, *14*, 4-17. doi.org/10.1007/s12024-017-9934-x.

Brodie, L. R., Odell, M., Ranson, D., Young, C., Kitching, F. and Ibrahim, J. E. (2019). Sudden natural death behind the wheel: Review of driver deaths and fitness to drive assessment history in Victoria, Australia 2012–2013. *Journal of Forensic and Legal Medicine*, *63*, 31-33. doi.org/10.1016/j.jflm.2019.02.011.

Euro NCAP (2022). *Euro NCAP Vision 2030: A Safer Future for Mobility*. https://cdn.euroncap.com/media/74468/euro-ncap-roadmap-vision-2030.pdf

Euro NCAP (2023). *Assessment Protocol – Safety Assist Safe Driving, v10.3*. https://www.euroncap.com/media/79883/euro-ncap-assessment-protocol-sa-safe-driving-v103.pdf

Fitzharris, M. P., Corben, B., Lenné, M. G., Pok Arundell, T., Peiris, S., Liu, S., Stephens, A., Fitzgerald, M., Judson, R., Bowman, D. M., Gabler, H. C., Morris, A. and Tingvall, C. (2020). *Overview and analysis of serious injury crashes – crash types, injury outcomes and contributing factors, ECIS Report 1*. Monash University Accident Research Centre.

Hanna, R. (2009). *The Contribution of Medical Conditions to Passenger Vehicle Crashes*. U.S. Department of Transportation, National Highway Traffic Safety Administration2.

Lindsay, V. L. A. and Baldock, M. R. J. (2008). Medical conditions as a contributing factor in crash causation. *Proceedings of the 2008 Australasian Road Safety Research, Policing and Education Conference,* 610-628

Marinella, M. A. (2004). A car crash complicating influenza. *Mayo Clinic Proceedings, 79*(8), 1087-1088.

Motozawa, Y., Yokoyama, T., Hitosugi, M. and Tokudome, S. (2005). Analysis of sudden natural deaths while driving with forensic autopsy findings. *19th International Technical Conference on the Enhanced Safety of Vehicles (ESV)*.

Neal, A., Carne, R., Odell, M., Ballek, D., D'Souza, W. J. and Cook, M. J. (2018). Characteristics of motor vehicle crashes associated with seizure: Car crash semiology. *Neurology, 91*(12), e1102-e1111. doi:10.1212/WNL.0000000000006208.

Sakurai, K., Yamamoto, J., Kurita, T., Takeda, Y. and Kusumi, I. (2014). Video event data recording of a taxi driver used for diagnosis of epilepsy. *Epilepsy and Behavior Case Reports, 2*, 24-25. dx.doi.org/10.1016/j.ebcr.2013.12.007.

Tervo, T. M. T., Neira, W., Kivioja, A., Sulander, P., Parkkari, K. And Holopainen, J. M. (2008). Observational Failures/Distraction and Disease Attack/Incapacity as Cause(s) of Fatal Road Crashes in Finland, *Traffic Injury Prevention, 9*(3), 211-216. doi.org/10.1080/15389580802040303.

# Examining driver fatigue during conditionally automated driving using contactless physiological measurement

Rory Coyne [1, 2*], Bilal Khan [3], Mohamed Moustafa [4], Ashkan Parsi[2], Waseem Shariff[2], Joseph Lemley[2], Alan F. Smeaton[5], Peter Corcoran[3], Jane C. Walsh[1].

[1] School of Psychology, University of Galway, Galway, Ireland
[2] Tobii Galway, Parkmore Industrial Estate Block 5, Galway, Ireland
[3] School of Psychology, Trinity College Dublin, Dublin, Ireland
[4] Department of Electrical and Electronic Engineering, University of Galway, Galway, Ireland
[5] School of Computing Dublin City University, Dublin, Ireland
[*]r.coyne7@universityofgalway.ie

**Abstract: As the level of vehicle automation increases, drivers can delegate greater responsibility for vehicle control to the automated system. During conditionally automated driving, while the driver can delegate full vehicle control to the system, they are expected to remain ready to regain control at all times. The monotonous nature of the driver's supervisory role during conditionally automated driving can result in elevated levels of fatigue. Thus, fatigue, which contributes to 10-20% of all road traffic accidents, poses a significant challenge to the safety of conditionally automated driving. One feature which will be required within all new vehicles from 2026 is a driver monitoring system, which can measure psychological states and intervene when such states reach unsafe levels. However, current monitoring technology relies upon invasive methods which are not suitable for the naturalistic driving context. The present work in progress seeks to measure fatigue during conditionally automated driving using contactless physiological measurement, through investigating the effect of prolonged automation on the development of driver fatigue.**

## 1. Introduction

Conditionally automated driving (also known as SAE Level 3 automation; SAE International, 2023) allows for the continuous performance of the driving task by an automated system within a specific set of environmental conditions referred to as the operational design domain (ODD). At this level of automation, the driver is required to remain ready to regain vehicle control. Any deviations from the ODD will prompt the automated system to issue a takeover request, in which the human operator is instructed to regain responsibility for the driving task within a short period of time.

Emergent driver monitoring systems (DMS) will soon be commonplace in all vehicles and will be responsible for making assessments about the driver's psychological state, thus assisting to oversee the safety of takeover requests by assessing the driver's readiness to regain control.

### *1.1 Driver fatigue*

Fatigue is reported to be a contributing factor in 10-20% of all road traffic accidents (Zhang et al., 2020). During conditionally automated driving, the driver's role changes from that of an active operator to a passive supervisor, and this can lead to elevated levels of fatigue. According to May & Baldwin's (2009) model of fatigue, passive fatigue occurs when low levels of workload are imposed by a task. Fatigued drivers also display low levels of situation awareness, which is important for "knowing what's going on so you can figure out what to do" (Adam, 2005, pp. 319). Previous research has demonstrated that fatigued drivers display poorer takeover performance during automated driving than drivers who are not experiencing fatigue (Feldhütter et al., 2019). There have been mixed findings for the effect of NDRT engagement on takeover performance. Recently, it has been suggested that NDRTs could be used to mitigate fatigue during automated driving by preventing mental underload. Nonetheless, there is ample evidence that the development of fatigue is a pressing safety issue during automated driving.

### *1.2 Measurement of psychological states during automated driving*

For a DMS to function in a naturalistic driving context, it will require contactless sensors to measure the driver's psychological state. Current methods of doing so rely on invasive, skin-contact sensors, and this has motivated the development of non-invasive alternatives. Infrared imaging has been proposed as one such solution. Near infrared (NIR) and far infrared (thermal) imaging can be used to measure changes in the temperature of the skin, which vary as a function of autonomic nervous system activity (Cardone et al., 2020). Infrared light penetrates further into human tissue than visible light due to its greater wavelength, thus allowing for visualisation of physiological changes below the skin's epidermis (Stemberger et al., 2010). Furthermore, infrared imaging does not depend on external illumination sources, thus making it a suitably robust technique for driver monitoring, in which illumination sources vary dramatically.

Infrared imaging has previously been applied to measure drowsiness (Tashakori et al., 2022) and cognitive load (Stemberger et al., 2010). However, infrared imaging has yet to be applied in an automated driving context. Furthermore, to date no research has attempted to measure fatigue using infrared imaging.

### *1.3 Objectives and research questions*

The present research seeks to investigate the effect of prolonged automation and NDRT performance on driver fatigue during conditionally automated driving, and to measure fatigue using contactless physiological measures. Specifically, the following research questions will be explored:

RQ1: What is the effect of prolonged automated driving on driver fatigue, as measured by drivers' physiological responses?

RQ2: What is the effect of non-driving-related task (NDRT) performance on driver fatigue, as measured by drivers' physiological responses?

RQ3: How accurately can physiological arousal be measured using infrared imaging, when compared to traditional physiological measures?

RQ4: How accurately can driver fatigue be classified based on infrared imaging-derived physiological data using machine learning methods?

RQ5: How does fatigue vary over the course of a prolonged automated drive?

## 2. Method

### *2.1 Design*

This driving simulator study will employ a within-subjects design. The simulated driving environment will consist of a daytime driving sequence on a three-lane highway.

The independent variable will be the experimental condition, with three levels (baseline, automated drive with NDRT, prolonged automated drive with no NDRT). Dependent variables will include physiological metrics (heart rate (HR), heart rate variability (HRV), breathing rate), infrared metrics (NIR-derived HR, HRV and breathing rate, skin temperature on the forehead, nose tip, and chin areas), eye-tracking (saccades - transient eye movements from one area of fixation to another) and self-reported fatigue and workload.

### *2.2 Participants*

Individuals 18 years or older, holding a full driving license, with no cardiovascular disease, no history of motion sickness, simulator sickness or vertigo, and normal or corrected-to-normal vision will be eligible to participate in this study. A total of 36 participants will be recruited.

### *2.3 Apparatus*

A high-fidelity driving simulator at the offices of Tobii Galway will be used. Drivers' physiological data will be collected using a Zephyr$^{TM}$ BioModule 3 at 250 Hz capable of measuring both electrocardiography (ECG) and respiration. Three cameras (NIR camera, FLIR thermal camera and an Event camera measuring driver saccades), connected to a desktop PC will be used to capture the infrared data.

### *2.4 Procedure*

Upon arrival to the laboratory, participants will be briefed on the details of the study and asked to provide their informed consent. Participants will then be provided with a practice drive session, which will involve interacting with the driving simulator and the NDRT. The NDRT in this study will be an auditory 2-back task, in which participants will be asked to respond verbally when the auditory stimulus matches the stimulus that was presented 2 stimuli previously. Following the practice drive, baseline measurements will be taken.

The two counterbalanced experimental conditions will be the NDRT condition and the prolonged automated drive condition. During the NDRT condition, participants will perform the 2-back task for 20 minutes during a simulated automated drive. In the prolonged automated drive condition, participants will monitor an automated drive for 2 hours, without performing an NDRT. During all three conditions, participants will wear the Zephyr$^{TM}$ BioModule sensor. At the end of each condition, participants will complete the NASA-TLX and Karolinska Sleepiness Scale (KSS) measures.

### *2.5 Analysis*

To address RQ1 and RQ2, Analysis of variance (ANOVA) will be performed to examine the effect of condition (baseline, automated drive with NDRT, prolonged automated drive with no NDRT) on drivers' physiological responses.

To address RQ3, relevant features will be extracted from the thermal image data. The coordinates of the three regions of interest will be determined, and the average value of the pixels of each region will be extracted. HR, HRV, and breathing rate will be extracted from the NIR signal using a convolutional neural network (CNN). The infrared data will then be correlated with the physiological measures and the self-reported fatigue data to determine the level of agreement between measures.

To address RQ4, supervised machine learning methods will be applied to the infrared data. Models will include support vector machine (SVM), random forest, Naïve Bayes, and decision trees. The mean absolute error (MAE) and root mean squared error (RMSE) will be used to evaluate performance.

To address RQ5, the physiological and infrared data will be plotted against time to obtain the temporal resolution of drivers' responses during the prolonged automated drive condition.

## 3. Conclusions

As the level of vehicle automation increases, drivers display increasing levels of fatigue, which can jeopardise the safety of transitions of control between automated system and driver. Current methods of measuring psychological states are dependent on invasive, skin-contact methods which are ill-suited for naturalistic driving. The present work in progress seeks to measure fatigue during conditionally automated driving using contactless physiological measurement. This work will inform the development of driver monitoring systems which will be responsible for assessing the driver's availability to respond to requests from an automated system to intervene. It will also have implications for the driver fatigue and workload literature, by investigating the effect of NDRT performance on driver fatigue, and how fatigue varies during prolonged automated driving.

## 4. Acknowledgments

This work is being conducted with the financial support of the Science Foundation Ireland Centre for Research Training in Digitally-Enhanced Reality (d-real) under Grant No. 18/CRT/6224. 

## References

Adam, E. C. (2005, October). Fighter cockpits of the future. In *[1993 Proceedings] AIAA/IEEE Digital Avionics Systems Conference* (pp. 318-323). IEEE.

Cardone, D., Perpetuini, D., Filippini, C., Spadolini, E., Mancini, L., Chiarelli, A. M., & Merla, A. (2020). Driver stress state evaluation by means of thermal imaging: A supervised machine learning approach based on ECG signal. *Applied Sciences, 10*(16), 5673.

Feldhütter, A., Ruhl, A., Feierle, A., & Bengler, K. (2019, October). The effect of fatigue on take-over performance in urgent situations in conditionally automated driving. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)* (pp. 1889-1894). IEEE.

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.

SAE International (2023). *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles.* Retrieved from: https://www.sae.org/standards/content/j3016/

Stemberger, J., Allison, R. S., & Schnell, T. (2010, May). Thermal imaging as a way to classify cognitive workload. In *2010 Canadian Conference on Computer and Robot Vision* (pp. 231-238). IEEE.

Tashakori, M., Nahvi, A., & Ebrahimian Hadi Kiashari, S. (2022). Driver drowsiness detection using facial thermal imaging in a driving simulator. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of engineering in medicine, 236*(1), 43-55.

Zhang, Q., Wu, C., & Zhang, H. (2020). Driving fatigue prediction model considering schedule and circadian rhythm. *Journal of advanced transportation, 2020*, 1-10.

# Can perceptual stimuli help reduce driver stress when regaining control after a monotonous episode in automated driving?

Laora Kerautret [1*], Antoine Favre-Felix [2], Stéphanie Dabic [1]

[1] Valeo Switches & Smart Controls Product Line, Rue Jules VERNE - Vetraz Monthoux
BP 509 - 74106 Annemasse Cedex, FRANCE
[2] INNOVATEAM – 52, Quai de Dion Bouton 92800 PUTEAUX
[*] laora.keraurtret@valeo.com

**Abstract: The monotonous episodes in automated driving are likely to lead to passive fatigue (PF), which in turn can impair driver alertness. Paradoxically, in automated driving, a good level of alertness is strongly required when taking back control (TOR) of the vehicle. Thus, drivers subjected to PF in automated driving could, at the TOR moment, perceive a low availability of attentional resources, which could subsequently trigger excessive stress. The aim of this study was therefore to examine the added value of perceptual stimuli in keeping the driver reasonably alert, and thus moderating TOR-induced stress. 46 drivers were exposed to PF during automated driving; half of them received various perceptual stimuli after 20 minutes of automated driving, while the other half received no stimuli at all. Preliminary results, based on perceived stress and heart rate, indicate that stress, experienced during TOR and after a monotonous episode, can be moderated by the use of perceptual stimuli during automated driving.**

## 1. Introduction

The monotonous environments are widely recognized to lead to states of cognitive underload, also known as passive fatigue (PF). Although PF is often inconsequential, it is particularly critical when it occurs in automated driving, where driver intervention may still be required. Indeed, one of the risks of PF in automated driving is that the driver, experiencing a decrease in alertness (Larue et al., 2011; Saxby et al., 2013), may perceive his/her available attentional resources as being inferior to those needed to regain control of the vehicle effectively and on time. Such a cognitive appraisal could then induce excessive stress in the driver, potentially leading to risky driving behavior and reduced performance (Ge et al., 2014; Rendon-Velez et al., 2016).

To avoid this domino effect, this study investigates the use of various perceptual stimuli (defined as stimuli or activities that appeal to the senses and are likely to capture attention) during automated driving in order to moderate PF, and thus maintain a certain threshold of alertness, which would be particularly beneficial in avoiding excessive stress when regaining control of the vehicle.

## 2. Method

### 2.1 Participants

46 drivers were recruited and distributed in equal proportions in the "stimuli" and "control" groups. All participants gave written consent.

### 2.2 Experimental design

The experiment took place on a real circuit with an automated vehicle. Before completing a training session, the participants were given these instructions: execute all requests from the automation system, i.e. either to delegate control or take back control, monitor the driving environment, and participate in the activities when offered by the system.

Then, participants were asked to drive manually for 5 minutes, then switch to automated mode for 40 minutes before taking over control (TOR) and driving manually again for 5 minutes. (**Fig. 1**).

Participants in the stimuli group were exposed to a 5-minute series of perceptual stimuli, starting from 20 minutes of automated driving and lasting for 20 minutes. This series included a general knowledge quiz, dynamic ambient lighting, guided mindfulness practice, and fresh air diffusion. The control group was not exposed to perceptual stimuli.

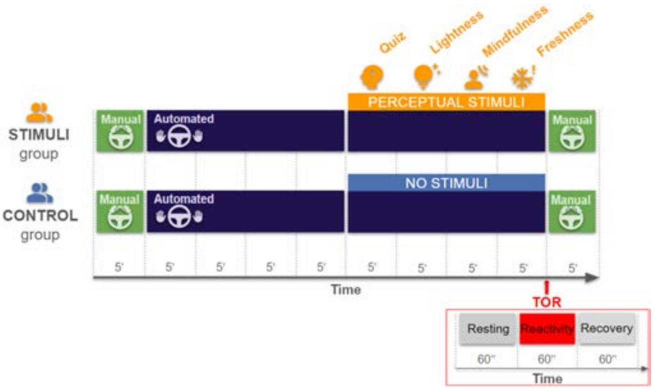



*Fig. 1. Experimental design.*

After 40 minutes of automated driving, participants were asked to *prepare* to take control of the vehicle, before *regaining* control 20 s later after passing a signal flag. TOR herein refers to these 2 stages, including preparation and action to regain control. TOR was explored over three 60-second time windows; *Resting* (from 60 s before asking participants to prepare for TOR), *Reactivity* (from the request to prepare for TOR to 60 s after the request) and *Recovery* (from 60s after asking participants to prepare for TOR).

### 2.3 Measurements

Perceived stress regarding TOR was assessed at the end of the experiment using a 5-point scale from 0 = “no stress” to 5 = “severe stress”.

Heart rate (HR) was measured throughout the experiment at a sampling rate of 500 Hz. Furthermore, to assess whether HR was modulated by perceived stress, the participants were divided into two groups: participants who reported a score greater than 0 on the perceived stress scale were included in "*Perceived stress*", while those who reported a score equal to zero were included in "*Non-perceived stress*" (**Table 1**).

***Table 1***. *Number of participants*

| | Group | | |
|---|---|---|---|
| | Control | Stimuli | Total |
| Non-perceived stress | 6 | 15 | **21** |
| Perceived stress | 17 | 8 | **25** |
| Total | **23** | **23** | **46** |

## 3. Results

### 3.1 Perceived stress

The mean perceived stress was 0.43 (SD = 0.66) for the stimuli group and 1.39 (SD = 1.34) for the control group. Paired *t*-tests indicated that perceived stress was significantly lower for the stimuli group compared to the control group ($t = 3.87$, $p < .001$, $d = 0.806$).

A chi-square test reported a significant relationship ($\chi^2_{(1, 46)} = 7.097$, $p < .01$) between group (control, stimuli) and subgroup (non-perceived stress, perceived stress). The stimuli group was less likely than the control group to report perceived stress at TOR time.

### 3.2 Heart rate

To investigate whether the stimuli group has an advantage in moderating TOR-related stress compared with the control group, a Linear Mixed Model (LMM) was performed using Group (stimuli, control) Perceived stress (perceived stress, non-perceived stress) and Time window (resting, reactivity, recovery) as fixed effects, and ID participant as a random effect (**Fig. 2**).

The LMM reported a significant main effect of time window ($F_{(2,\ 15985)} = 988.00$, $p < .001$). Contrast analyses (Bonferroni correction) indicated a higher HR for *Reactivity* compared to *Resting* ($z = 37.31$, $p < .001$) and *Recovery* ($z = 39.548$, $p < .001$), as well as lower HR for *Recovery* compared to *Resting* ($z = -2.828$, $p < .05$).

The LMM reported a significant interaction between time window and perceived stress ($F_{(2,\ 15985)} = 130.60$, $p < .001$). Contrast analyses (Bonferroni correction) showed a higher HR for the *Perceived stress group* compared to the *Non-perceived stress* group during *Reactivity* ($z = 4.25$, $p < .001$) and Recovery ($z = 2.49$, $p < .05$), but no HR change during *Resting* ($z = -0.006$, $p > .05$).

The LMM reported a significant interaction between time window, perceived stress and group ($F_{(2,\ 15985)} = 77.29$, $p < .001$). Contrast analyses (Bonferroni correction) indicated within the *Perceived stress* group, no change in HR between the control and stimuli groups for *Resting* ($z = -0.015$, $p > .05$), *Reactivity* ($z = 0.839$, $p > .05$), and *Recovery* ($z = -0.235$, $p > .05$). However, within the *No perceived stress* group, a lower HR was observed in *Reactivity* for the stimuli group *versus* control group ($z = -2.824$, $p < .05$), but any HR change was reported between the stimuli and control groups regarding *Resting* ($z = -0.006$, $p > .05$) and *Recovery* ($z = 0.300$, $p > .05$).

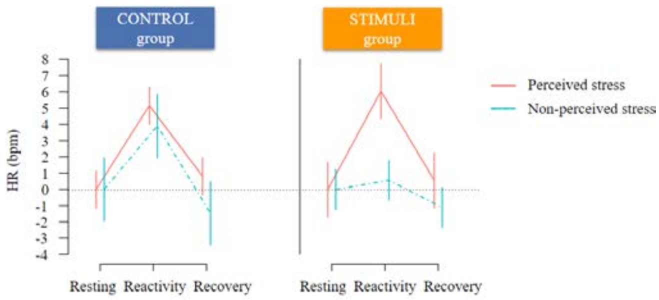



***Fig. 2.*** *HR responses.*

## 4. Discussion

The aim of the study was to determine whether perceptual stimuli could have a safety benefit for drivers, by reducing the stress experienced when regaining control of the vehicle after a monotonous episode in automated driving.

The results indicate that providing perceptual stimuli during automated driving reduced the overall level of perceived stress during TOR. The overall level of reduced perceived stress may be explained in part by the fact that only a third of participants (N = 8) in the stimuli group reported feeling stress, compared with two-thirds (N = 17) in the control group, suggesting that perceptual stimuli may contribute to stress-free perception of TOR.

The results reveal that overall, an increase in HR was observed in both groups from the moment the system asked participants to prepare to resume control (i.e. during *Reactivity*) and beyond control regaining (i.e. during *Recovery*). Similar increase in HR was previously observed in driving for various road hazards (Johnson et al., 2011; Schmidt-Daffy, 2013) including emergency TOR (Kerautret et al., 2023a), and was interpreted as a stress response akin to flight defensive behaviour (Keraautret et al., 2023b).

The outcomes also show that HR after the request to TOR preparation (i.e. during *Reactivity* and *Recovery*) was greater for the participants who declared perceived stress, compared to those who reported non-perceived stress. This result is similar to that of a previous study which showed that perceived stress modulated the amplitude of HR response after exposure to road hazards (Keraautret et al., 2022).

Finally, participants (N = 15) exposed to perceptual stimuli and claiming not to have perceived stress showed a lower HR increase during *Reactivity*, compared to participants (N = 6) also claiming not to have perceived stress but not being exposed to perceptual stimuli. This result, combined with the lower number of participants who perceived stress in the stimuli group, suggests that perceptual stimuli presented a first benefit in reducing the likelihood of experiencing stress during TOR preparation, and a second benefit in reducing the increase in HR in the majority of participants.

## 5. Conclusions

These preliminary results encourage the use of perceptual stimuli during automated driving episodes likely to induce PF. Indeed, by counteracting the deleterious PF effects (i.e. excessive loss of alertness), certain perceptual stimuli could have the advantage of preserving attentional resources and thus preventing excessive stress when it is required to regain control of the vehicle.

## 6. Acknowledgments

This research was undertaken with the financial support of the BPI as part of a collaborative project entitled SERENITE. Sabine Lewkowicz also made a significant contribution to the project, creating and lending her voice to the recordings dedicated to guided mindfulness practice.

## References

Ge, Y., Qu, W., Jiang, C., Du, F., Sun, X., & Zhang, K. (2014). The effect of stress and personality on dangerous driving behavior among Chinese drivers. Accident Analysis & Prevention, 73, 34–40. doi: 10.1016/j.aap.2014.07.024.

Johnson, M. J., Chahal, T., Stinchcombe, A., Mullen, N., Weaver, B., & Bédard, M. (2011). Physiological responses to simulated and on-road driving. International Journal of Psychophysiology, 81(3), 203–208. doi: 10.1016/j.ijpsycho.2011.06.012.

Kerautret, L., Dabic, S., & Navarro, J. (2023a). Exploration of driver stress when resuming control from highly automated driving in an emergency situation. Transportation research part F: traffic psychology and behaviour, 93, 222-234. doi: 10.1016/j.trf.2023.01.016

Kerautret, L., Dabic, S., & Navarro, J. (2023b). Exploring hazard anticipation and stress while driving in light of defensive behavior theory. Scientific reports, 13(1), 7883. doi: 10.1038/s41598-023-34714-7.

Kerautret, L., Dabic, S., & Navarro, J. (2022). Detecting driver stress and hazard anticipation using real‐time cardiac measurement: A simulator study. Brain and behavior, 12(2), e2424. doi: 10.1002/brb3.2424.

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. doi: 10.1016/j.aap.2011.05.023

Rendon-Velez, E., Van Leeuwen, P. M., Happee, R., Horvath, I., van der Vegte, W. F., & de Winter, J. C. (2016). The effects of time pressure on driver performance and physiological activity: A driving simulator study. Transportation research part F: Traffic psychology and behaviour, 41, 150–169. doi: 10.1016/j.trf.2016.06.013

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. doi: 10.1037/a0034386

Schmidt-Daffy, M. (2013). Fear and anxiety while driving: Differential impact of task demands, speed and motivation. Transportation Research Part F: Traffic Psychology and Behaviour, 16, 14–28. doi: 10.1016/j.trf.2012.07.002.

# Impact of Tesla Autopilot Driving on Drivers' Cognitive Workload and Glance Allocation

Francesco N. Biondi, Ph.D. [1], Noor Jajo[2]

[1] Human Systems Lab, Department of Kinesiology, University of Windsor, 2555 College Ave, Windsor, ON, Canada
[2] Department of Psychology, University of Utah, Salt Lake City, UT, US
*francesco.biondi@uwindsor.ca

**The use of partially automated or SAE L2 vehicles is expected to change the role of the human driver from operator to supervisor, which may have an effect on the driver's cognitive workload and visual attention. In this study, 30 Ontario drivers operated a 2022 Tesla Model 3 in manual and Autopilot/L2 mode. Cognitive workload was measured by means of the Detection Response Task, and visual attention was measured by means of coding glances on and off the forward roadway. No difference in cognitive workload was found between driving modes. However, during L2 driving, drivers spent less time glancing at the forward roadway, and more time glancing at the vehicle's touchscreen. These data add to our knowledge of how vehicle automation affects cognitive workload and attention allocation. They also show potential safety risks that the adoption of partially automated driving may have on driver distraction.**

## 1. Introduction

The Society of Automotive Engineers defines six levels of automated driving systems from fully-manual to fully-automated (SAE, 2021). A level 2 system maintains control of the vehicle's longitudinal and lateral behaviour and the human driver is responsible for actively monitoring its functioning and resuming manual control whenever necessary. The adoption of these systems comes with intended safety benefits. Yet, a still limited yet growing body of research has shown some potential safety risks of operating L2 systems. The Human Factors literature posits changes in drivers' cognitive workload as a potential risk associated with operating L2 systems. When in L2 mode, the role of the human driver will switch from that of vehicle operator to that of vehicle supervisor (Biondi et al., 2019; Cabrall et al., 2019). We hypothesize that such drastic transition in the driver's responsibilities will result in a decline in cognitive workload that is accompanied by a reduction in their ability to effectively monitor the functioning of the L2 system (Merat et al., 2019; Strayer et al., 2020). Consistent with this hypothesis, prior work observed a reduction in cognitive workload when the L2 system was operational (Biondi et al., 2018; Heikoop et al., 2019). Conflicting findings come from the studies by Strayer and colleagues (Lohani et al., 2021; McDonnell et al., 2021) wherein no differences in cognitive workload were observed between manual and L2 driving.

Changes in cognitive workload are expected to influence the driver's ability to maintain visual attention toward the driving task when the L2 system is operational. It is hypothesized that drivers might seek engagement in non-driving activities in conditions of boredom as a way to counter the declining cognitive workload – a phenomenon known as proactive self-regulation (Strayer & Fisher, 2016). With this, it is expected that they will spend less time attending to the driving task and more time allocating attention to driving-unrelated activities. Work by Gaspar and Carney (2019) shows that, when the L2 system is operational, drivers spend more time glancing at the in-vehicle touchscreen (see Gershon et al., 2023, and Noble et al. 2021 for similar results).

With this in mind, this study's objectives are: 1) investigating the effect that L2 driving has on cognitive workload, and 2) exploring the relationship between cognitive workload and drivers' visual attention.

## 2. Method

### *2.1 Participants*

30 volunteers (13 females) participated in this study. Their average age was 22 years old and standard deviation of age was 4.36 years. A University of Windsor Research Ethics Board approval (REB #20-141) was obtained for the study.

### *2.2 Design*

A factorial design with one independent factor: driving mode, was adopted in this study. Participants drove the vehicle in one of two modes: manual or L2. Dependent measures included: performance to the ISO DRT; total eyes off the road time (TEORT). Other metrics including the total glance time by area of interest (AOI), and average and maximum glance duration by AOI were recorded and analysed but are not included here due to the word restriction.

### *2.3 Equipment*

**Vehicle**. A 2022 Tesla Model 3 was used for the study, which can be driven in either manual, L1 or L2 mode. Our analysis will focus on the comparison between manual and L2 driving. **DRT**. The vibrotactile version of the ISO DRT was used in the study. Upon the presentation of a vibrotactile stimulus which occurred every 3–5 s, participants were instructed to press the microswitch as fast as possible. Reaction times (RT) in milliseconds were recorded. **Cameras**. The vehicle was retrofitted with three GoPro cameras that offered views of the driver, the forward roadway, and the touchscreen. Four areas of interest (AOI) were identified: forward roadway, side mirrors, rearview mirror, and touchscreen. Each drive was analysed by two coders and any discrepancies in the coding were reviewed for consistency by a third coder. A Cohen's kappa of 0.85 was found indicating very strong consistency agreement. Total eyes off the road time (TEORT) was calculated as the summation of all glance durations to all AOIs other than the forward roadway during a sample interval in seconds. This represents an established metric for measuring the safety of secondary tasks (Monk et al., 2023; NHTSA, 2013). Total glance time by AOI, average and maximum glance durations by AOI were also calculated but are not presented here due to the word limit.

### *2.4 Procedure and analysis*

Participants drove on the section of Ontario highway 401 between Windsor and Chatham twice: one in manual mode and one in L2 mode. The order of the two drives was counterbalanced. Each drive had a duration of 40 minutes and was kept consistent across participants. Participants were instructed to take exit 81 in Chatham and park at a gas station where they could take a 15-minute break. After the break, they re-entered the highway in the opposite direction and the second drive begun. The second drive ended at exit 13 in Windsor at which point participants were instructed to drive back to the University of Windsor campus. The experimental phase took up to 2 hours. Bayes factor analyses were conducted to investigate the effects of the factor driving mode on the dependent measures (Held & Ott, 2018; Quintana & Williams, 2018).

## 3. Results

### *3.1 Investigate the effect that L2 driving has on cognitive workload.*

DRT performance was analysed for this objective. A Bayesian t-test with mode (2 levels: manual, L2) was conducted to investigate differences between the two modes. A BF of 0.33 was found indicating that no differences in average RT were found between manual (M = 541 milliseconds) and L2 mode (M = 551 milliseconds).

### *3.2 Explore the relationship between cognitive workload and drivers' visual attention.*

TEORT, total glance time, average and maximum glance durations were analysed for this objective, but only TEORT is presented here. Figure 1 shows TEORT in manual and L2 mode. A Bayesian t-test was conducted to investigate the effect of mode on TEORT. A BF of 180.53 was found indicating strong evidence that TEORT increased when drivers operated the vehicle in L2 mode. Whereas TEORT averaged approximately 100 seconds or approximately 4% of the entire drive time in manual mode, it increased to 200 seconds on average of approximately 8% of the entire drive time in L2 mode.

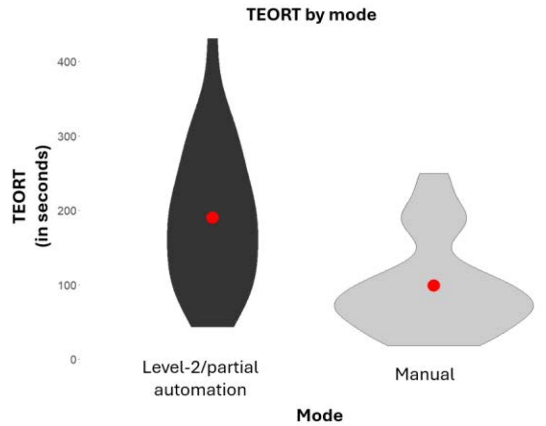



*Figure 1. Violin plot showing TEORT distributions by mode. The red dots represent average TEORT in the two modes: SAE level-2 partial automation, and manual driving.*

## 4. Discussion

This study's first objective was to investigate the effect that L2 driving had on cognitive workload. Analyses conducted on DRT RT and accuracy revealed Bayes factors of 0.33 and 0.21, respectively. Adopting Lee and Wagenmakers (2013)'s interpretative model, these indicate moderate evidence in support of the null hypothesis that no difference in cognitive workload were found between manual and L2 driving. This pattern is consistent with prior work by Lohani et al. (2021) where drivers' cognitive workload was not affected by driving mode. When the L2 system is operational, drivers are no longer in charge of steering or controlling the accelerator or brake pedals. Human Factors literature on vehicle automation suggest that this may result in a reduction of the driver's cognitive workload, a hypothesis that appears not to be supported by our data.

This study's second objective was to further investigate the effect of L2 driving on drivers' glance allocation, and its relationship with cognitive workload. Analyses revealed an increase in the total eyes of the road time in L2 mode with participants glancing away from the forward roadway an average of 200 seconds or 8% of the entire drive time in L2 mode and 100 seconds or 4% of the entire drive time in manual mode. Looking at figure 1 it is also interesting to note the wider distribution in L2 mode suggesting that, whereas TEORT was relatively homogenous during manual driving, it became more variable when the L2 system was operational with some participant's TEORT well exceeding 300 seconds or approximately 12% of the entire drive time. This pattern finds support in the work by Gaspar and Carney (2019) and Noble et al. (2021) that also found an increase in TEORT during L2 driving.

Our study has limitations. First, participants were not experienced L2 system users, which may limit the applicability of our findings. It is also worth noting that, although no difference in DRT performance was found between manual and L2 driving, the two conditions may still have required distinct levels of driving demand.

## 5. Conclusions

Findings for objectives 1 and 2 show that, while drivers' cognitive workload appeared to be unaffected by driving in L2 mode, drivers spent more time looking away from the road when the Autopilot system was engaged. When combined with the existing literature on partial automation, we speculate that, when the L2 system is engaged, drivers may be more inclined to look away from the road to counteract boredom. It follows that the risk of driver inattention toward safety-relevant events in the driving scene may be greater when the L2 system is active.

## 6. Acknowledgments

We acknowledge the generous contribution by NSERC, SSHRC, and the Ontario Ministry of Transportation.

## References

Biondi, F., Alvarez, I., Jeong, K., Biondi, F., Alvarez, I., & Jeong, K. (2019). Human - System Cooperation in Automated Driving. *International Journal of Human–Computer Interaction*, *00*(00), 1–2. https://doi.org/10.1080/10447318.2018.1561793

Biondi, F. N., Lohani, M., Hopman, R., Mills, S., Cooper, J. M., & Strayer, D. L. (2018). 80 MPH and out-of-the-loop : Effects of real-world semi-automated driving on driver workload and arousal . *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1878–1882. https://doi.org/https://doi.org/10.1177/1541931218621427

Cabrall, C. D. D., Eriksson, A., Dreger, F., Happee, R., & Winter, J. De. (2019). Theoretical Issues in Ergonomics Science How to keep drivers engaged while supervising driving automation ? A literature survey and categorisation of six solution areas. *Theoretical Issues in Ergonomics Science*, *0*(0), 1–38. https://doi.org/10.1080/1463922X.2018.1528484

Gaspar, J., & Carney, C. (2019). The Effect of Partial Automation on Driver Attention: A Naturalistic Driving Study. *Human Factors*, *61(8)*, 1261-1276. https://doi.org/10.1177/0018720819836310

Gershon, P., Mehler, B., & Reimer, B. (2023). Driver response and recovery following automation initiated disengagement in real-world hands-free driving. *Traffic Injury Prevention*, *24*(4), 356–361. https://doi.org/10.1080/15389588.2023.2189990

Heikoop, 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/10.1016/j.trf.2019.07.024

Held, L., & Ott, M. (2018). On p-Values and Bayes Factors. *Annual Review of Statistics and Its Application*, *5*, 393–419. https://doi.org/10.1146/annurev-statistics-031017-100307

Lee, M. D., & Wagenmakers, E. J. (2013). Bayesian cognitive modeling: A practical course. *Bayesian Cognitive Modeling: A Practical Course*, 1–264. https://doi.org/10.1017/CBO9781139087759

Lohani, M., Cooper, J. M., Erickson, G. G., Simmons, T. G., McDonnell, A. S., Carriero, A. E., Crabtree, K. W., & Strayer, D. L. (2021). No Difference in Arousal or Cognitive Demands Between Manual and Partially Automated Driving: A Multi-Method On-Road Study. *Frontiers in Neuroscience*, *15*(June), 1–12. https://doi.org/10.3389/fnins.2021.577418

McDonnell, A. S., Simmons, T. G., Erickson, G. G., Lohani, M., Cooper, J. M., & Strayer, D. L. (2021). This Is Your Brain on Autopilot: Neural Indices of Driver Workload and Engagement During Partial Vehicle Automation. *Human Factors*. https://doi.org/10.1177/00187208211039091

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), 87–98. https://doi.org/10.1007/s10111-018-0525-8

Monk, C., Sall, R., Lester, B. D., & Higgins, J. S. (2023). Visual and cognitive demands of manual and Voice-based driving mode implementations on smartphones. *Accident Analysis and Prevention*, *187*(January 2022), 107033. https://doi.org/10.1016/j.aap.2023.107033

NHTSA. (2013). *Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices* (Vol. 78, Issue 81).

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 and Prevention*, *151*(March 2020), 105959. https://doi.org/10.1016/j.aap.2020.105959

Quintana, D. S., & Williams, D. R. (2018). Bayesian alternatives for common null-hypothesis significance tests in psychiatry: A non-technical guide using JASP. *BMC Psychiatry*, *18*(1), 1–8. https://doi.org/10.1186/s12888-018-1761-4

SAE. (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. https://www.sae.org/standards/content/j3016_202104/

Strayer, D. L., & Fisher, D. L. (2016). SPIDER : A Framework for Understanding Driver Distraction. *Human Factors*, *58*(1), 1–12. https://doi.org/10.1177/0018720815619074

Strayer, D. L., Getty, D., Biondi, F., & Cooper, J. M. (2020). The Multitasking Motorist and the Attention Economy. In S. M. Lane & P. Atchley (Eds.), *Human Capacity in the Attention Economy*. APA Press.

# Eye, steering, and hands on wheel behaviors indicating driver engagement in assisted driving

Emma Tivesten [1*,2], Thomas Streubel [1], Mikael Ljung Aust[1]

[1] Volvo Cars Safety Centre, Volvo Cars, Göteborg, Sweden
[2] Chalmers University of Technology, Göteborg, Sweden
[*]emma.tivesten@volvocars.com

**Abstract: Assisted driving with infrequent need for driver input can lead to driver disengagement. A test track study was analyzed to investigate whether drivers' eye, steering, and hands on wheel behavior during uneventful assisted driving could be used to classify drivers' response to a conflict at the end of the drive. Long off-path glances, no driver steering input, and hands off wheel all increased the risk of a crash/near-crash. On the other hand, highly active steering, and large variation in lane position indicated of driver engagement, even for participants with many long off-path glances. To reliably assess driver disengagement during assisted driving, it is highly beneficial to combine eye, steering and hands on wheel behaviors.**

## 1. Introduction

Assisted driving systems (level 2; SAE International 2021) are evolving towards highly reliable performance for extended periods with little or no need for driver input. This leads to passive supervising which could result in reduced vigilance, mind wandering, increased secondary task engagement, and reduced eyes on path (Dunn et al, 2021, Körber et al, 2015; Morando et al, 2021), also referred to as driver disengagement (Lee et al., 2014). Introducing hands off driving enables manual secondary tasks, removes system feedback through hand to steering wheel contact (Mueller et al, 2021), and delays response to lateral control events (Larsson et al, 2022; Garbacik et al, 2021).

Several behaviors can be observed to assess driver engagement, including steering wheel torque input, hands on/off steering wheel, and visual behavior. In a previous test track study, 28% of the participants crashed with a conflict object after 30 minutes of highly reliable supervised assisted driving (Victor et al, 2018). Here, almost all crashers either had low levels of eyes on path, long visual response times to attention reminders, or displayed gaze concentration to the forward path during the drive (Tivesten et al, 2019). Interestingly, whether they had their hands on the wheel or not did not influence their conflict response (Pipkorn et al, 2021).

Streubel et al (2024), replicated the test setup in Victor et al (2018), using a level 2 system that required much lower steering wheel torque input to override lane centering compared to Victor et al (2018). Information about system capabilities and whether hands-off was allowed or not was varied between participants. Overall, 22 % of the participants had a crash or a near-crash. Crashes/near-crashes happened in all test conditions.

In this paper, we further analyze the data from this test with the aim to investigate whether metrics related to eye, steering and hands on wheel behaviors during uneventful driving can be used to classify whether participants are likely to crash/near-crash or not.

## 2. Method

### *2.1 Procedure and dataset*

The dataset includes 54 participants from a test track study described in Streubel et al (2024) with available data on eye, steering, and hands on/off wheel. The participants experienced 30 minutes of uneventful assisted car following, completing 5 laps on a test track. At the end of the drive, the lead vehicle cut-out and revealed a balloon car (approx. 3 seconds time to collision). The participants needed to steer to avoid a crash.

A development level 2 system was used with lane centering and distance keeping characteristics similar to an in-production Pilot Assist system. The system did not provide any steering or attention reminders. All participants were instructed to supervise the driving.

A binary variable categorized participants according to conflict event severity: Crash/near-crash (yes/no). The data included vehicle signals (i.e., steering wheel torque, lateral position in lane) and manually coded time series from video (i.e., eyes on path, hands on wheel) from the complete drive. Four metrics were derived and included in the analysis:

*GD2:* Number of off path glances longer than 2 seconds per hour of driving (N/h)
*ActTQ:* Percentage of time with active driver steering torque input (%)
*HoW:* Percentage of time with hands on wheel (%)
*SDLP:* Standard deviation on lateral position (m)

The GD2 metric included data from the complete drive since long glances are rare for some participants. The remaining metrics, all addressing lateral control, included data from the last lap only (i.e., the last 6 minutes prior to the conflict). See Appendix A for details.

*2.2 Analysis*

Receiver operating characteristics (ROC) curves were plotted for each metric. One or a few thresholds were graphically selected for each metric using the ROC-curves. The performance of each indicator (i.e., a metric and a threshold) was determined based on the True Positive Rate (TPR) and False Positive Rate (FPR) in correctly classifying participants according to the categories crash/near-crash: yes/no.

First, all pairs of specific indicators (i.e., low FPR) were combined using OR statements. Then, these pairs were combined with the most sensitive indicators using an AND statement. Two combinations with the highest performance (i.e., high TPR and low FPR) are presented in the results.

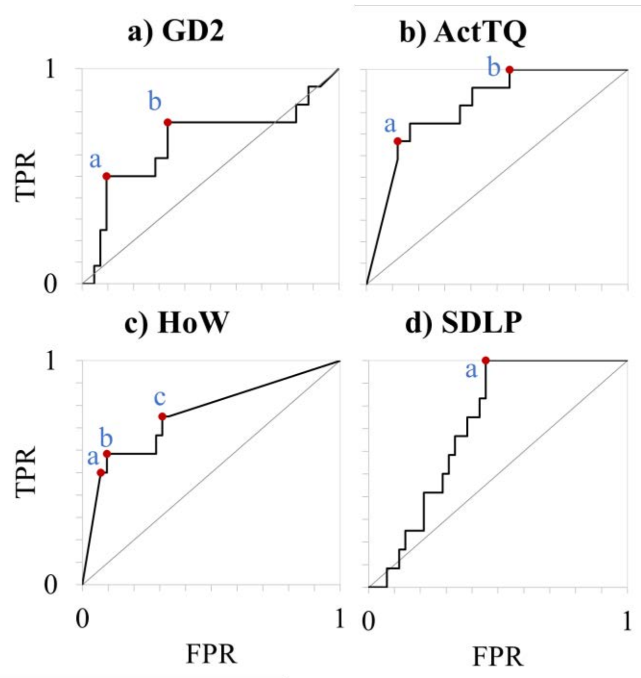



***Fig 1.*** *ROC-curves for the four metrics. The metric name is indicated in each subplot and the selected thresholds are highlighted by red markers and blue letters.*

## 3. Results

The crashers/near-crashers generally have more long off-path glances (GD2), less time with active driver steering input (ActTQ), lower hands on wheel time (HoW), and lower SDLP compared to the participants that did not have a crash/near-crash (Table 1).

***Table 1:*** *Descriptive statistics for each metric for the categories Crash/Near-crash (C/NC): no, yes. The left column shows the metric name and the Spearman correlation with the C/NC variable.*

| **Metric** $r_s$ *p* | | C/NC | N | Mdn | M | SD | Min | Max |
|---|---|---|---|---|---|---|---|---|
| GD2 *0,227* *p = .098* | N/h | no | 42 | 12,6 | 21,0 | 25,0 | 0,0 | 108,6 |
| | | yes | 12 | 33,8 | 38,6 | 30,9 | 0,0 | 88,6 |
| ActTQ *-.481* *p < 0.001* | % | no | 42 | 53,5 | 50,2 | 34,8 | 0,0 | 95,3 |
| | | yes | 12 | 0,0 | 12,1 | 21,3 | 0,0 | 61,4 |
| HoW *-.414* *p = .002* | % | no | 42 | 100,0 | 83,0 | 35,0 | 0,0 | 100,0 |
| | | yes | 12 | 1,9 | 41,6 | 50,8 | 0,0 | 100,0 |
| SDLP *-.311* *p= .022* | m | no | 42 | 0,087 | 0,106 | 0,056 | 0,056 | 0,365 |
| | | yes | 12 | 0,069 | 0,068 | 0,005 | 0,059 | 0,075 |

Table 2 shows that the most specific indicators (i.e., low FPR) are having little or no hands on wheel (HoWa,b), a high number of long off-path glances (GD2a), close to no active steering (ActTQa), followed by the lower threshold for long off-path glances (GD2b). On the other hand, low to moderate SDLP (SDLPa), and actively steering less than 62% of the time (ActTQb) are highly sensitive (TPR = 1) and less specific indicators since they include all crash/near-crash participants and approximately half of the remaining participants.

***Table 2:*** *Performance of each metric, and threshold in terms of FPR, TPR, and accuracy. Each threshold has an assigned index (a, b, c) presented next to the metric name. Area Under the Curve (AUC) indicate the overall metric performance in the left column.*

| **Metric (AUC)** | | **Threshold** | **Unit** | **FPR** | **TPR** | **Accuracy** |
|---|---|---|---|---|---|---|
| GD2 | a | ≥ 44,31 | N/h | 0,095 | 0,500 | 0,815 |
| (0,658) | b | ≥ 18 | | 0,357 | 0,750 | 0,667 |
| ActTQ | a | ≤ 0,022 | % | 0,119 | 0,667 | 0,833 |
| (0,832) | b | ≤ 62 | | 0,548 | 1,000 | 0,574 |
| HoW | a | = 0 | % | 0,071 | 0,500 | 0,833 |
| (0,758) | b | ≤ 3,9 | | 0,095 | 0,583 | 0,833 |
| | c | ≤ 99,1 | | 0,310 | 0,750 | 0,704 |
| SDLP (0,724) | a | ≤ 0,075 | m | 0,452 | 1,000 | 0,648 |

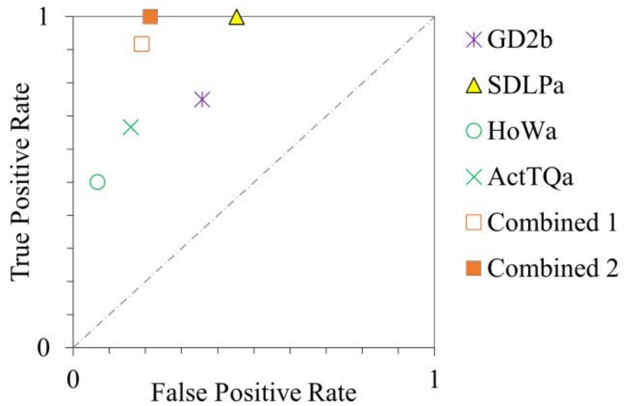



***Fig. 2.*** *Classification performance of crash/near-crash for single and combined indicators.*

Fig. 2 shows the two combinations of indicators with high performance, and the individual indicators included in these. The logic for these combinations is presented in Table 3.

For instance, Combined 1 classify a participant as positive (i.e., a likely crash/near-crash outcome) if they have both low to moderate SDLP (SDLPa), combined with either many long off-path glances (GD2b), or no hands on wheel (HoWa), or both (GD2b and HoWa). In combined 2, almost no active steering (ActTQa) is simply replacing no hands on wheel.

***Table 3:*** *The definition and performance of the combined metrics and thresholds*

| Combinations | FPR | TPR | Accuracy |
|---|---|---|---|
| Combined 1<br>(GD2b OR HoWa) AND SDLPa | 0,190 | 0,917 | 0,833 |
| Combined 2<br>(GD2b OR ActTQa) AND SDLPa | 0,214 | 1,000 | 0,833 |

## 4. Discussion

Overall, the results suggest that it is advantageous to combine visual and lateral control indicators to assess driver disengagement. This is a different outcome compared to (Tivesten et al, 2019; Victor et al, 2018), which found that visual behavior alone indicated driver disengagement. A possible explanation for this difference is that the drivers in the present study could override the vehicle lane centering force at a much lower steering wheel torque, compared to the previous study. Consequently, most drivers actively steered during the drive, and this seems to have secured driver engagement. However, if active driver steering requires too much effort it may cease, which means metrics of visual attention would remain the only way to assess driver engagement. Off-path glances longer than 2 seconds were more common for the crash/near-crash participants, a result that is consistent with Tivesten et al (2019). In addition, periods without driver steering and hands off wheel seems to increase the risk of a crash/near-crash. Even when visually attentive, one participant crashed after driving hands off. Designing the functions so that hands on wheel remains a natural and intuitive driver behavior also during assisted driving thus seems important, both to maintain driver engagement and to be able to detect disengagement.

Further research is needed to understand whether these findings generalize to other driver populations and contexts. Additional experiments could investigate how altered feedback strategies, such as different time intervals for hands on wheel or steering reminders, might influence driver engagement. Future analysis will also include additional metrics and analysis methods to further explore this topic.

## 5. Conclusions

This study found several indicators of driver engagement and disengagement in assisted driving. A high percentage of active driver steering input and high SDLP seems to secure driver engagement, while periods without driver steering or hands off wheel indicate driver disengagement. In addition, when there were no or moderate driver steering, long off-path glances indicated driver disengagement. It follows that eye, steering, and hands on wheel metrics needs to be combined to assess driver engagement.

## 6. Acknowledgments

The authors would like to thank their colleagues Ruben Brau, Patrik Andersson, Mirta Zelenika Zeba, and Emma Nilsson for supporting the study. Furthermore, we appreciate the annotation effort made by Lovisa Högberg, Jennifer Nordqvist and Rickard Nilsson.

## References

Dunn, N., Dingus, T., Soccolich, S., & Horrey, W. (2021). Investigating the impact of driving automation systems on distracted driving behaviors. *Accident Analysis & Prevention*, 156. https://doi.org/10.1016/j.aap.2021.106152.

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

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

Larsson, A., Sander, U., & Seppelt, B. (2022). To have and to hold during L2: Hands on wheel keeps drivers in lane while mind on road speeds response. *Proceedings of 8th International conference on driver distraction and inattention, DDI*. https://ddi2022.org/wp-content/uploads/2022/10/DDI2022-ABSTRACTS.pdf

Lee, J. D. (2014). Dynamics of driver distraction: The process of engaging and disengaging. *Association for Advancement of Automotive Medicine*, 58, 24–32

Morando, A., Gershon, P., Mehler, B., & Reimer, B. (2021). Visual attention and steering wheel control: From engagement to disengagement of Tesla Autopilot. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 65(1), 1390-1394. https://doi.org/10.1177/1071181321651118

Mueller, A. S., Reagan, I. J., & Cicchino, J. B. (2021). Addressing Driver Disengagement and Proper System Use: Human Factors Recommendations for Level 2 Driving Automation Design. In *Journal of Cognitive Engineering and Decision Making* (Vol. 15, Issue 1).

Pipkorn, L., Victor, T. W., Dozza, M., & Tivesten, E. (2021). Driver conflict response during supervised automation: Do hands on wheel matter? *Transportation Research Part F: Traffic Psychology and Behaviour*, 76, 14–25. https://doi.org/10.1016/j.trf.2020.10.001

SAE International. (2021). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles (SAE Standard J3016, Report No. J3016-202104). https://www.sae.org/standards/content/j3016_202104/

Streubel T., Tivesten E., Zeba M.Z., (2024). Conflict response in assisted driving with hands on or off wheel and different steering wheel torque settings. Manuscript submitted for publication.

Tivesten, E., Victor, T. W., Gustavsson, P., Johansson, J., & Aust, M. L. (2019). Out-of-the-loop crash prediction: The automation expectation mismatch (AEM) algorithm. *IET Intelligent Transport Systems*, 13(8). https://doi.org/10.1049/iet-its.2018.5555

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.

**Appendix A: Definitions of metrics, thresholds, and categories.**

| Heading | Category | Description |
|---|---|---|
| **Video (@30 Hz)** | | |
| | Driver face/upper body | |
| | Steering wheel | |
| | Forward roadway | |
| **Vehicle Signals (@100 Hz)** | | |
| | Steering wheel torque (Nm) | Signal from steering wheel torque sensor that mainly measure the torque that the driver applies to the steering wheel. |
| | Lateral position (m) | Measures lateral distance between the midline of the car and the center of the travel lane. |
| **Time series based on manual coding of videos from complete drive (@30 Hz)** | | |
| | Eyes on path | Binary time series for eyes on (1) versus eyes off path (0) |
| | Hands on wheel | Binary time series for hands on (1) versus hands off steering wheel (0). Hands on include any hand to steering wheel contact. |
| **Metrics – aggregated during the complete drive or last lap (i.e., last 6 minutes excluding the conflict)** | | |
| | GD2 | The number of off-path glances longer than 2 seconds per hour of driving during the complete drive [N/h] |
| | ActTQ | The percentage of time with driver active steering during last lap [%]. Active driver steering is defined as steering wheel torque outside a corridor [-0.4;0.25 Nm]. All hands-off driving were fully within this corridor. |
| | HoW | The percentage of time with hands on wheel during last lap [%]. |
| | SDLP | Standard deviation of the lateral position during last lap [m] |
| **Crash/near-crash (C/NC): yes/no** | | |
| | yes | This category included 12 participants that either had a crash or a near-crash during the conflict at the end of the drive.<br>In total, 6 participants had a crash (i.e., car-object impact), and 6 participants had a near-crashes. Near-crashes included a lateral distance to the balloon car less than 0.5m, or steering onset less than 1.1s TTC (time to collision) based on cluster analysis presented in Streubel et al (2024). |
| | no | This category includes the remaining 42 participants that did not have a crash or near-crash. |
| **Participants** | | |
| | Inclusion criteria | Volvo cars employees with a valid Swedish driver license. At least 5000 km driving during the last year. |
| | Exclusions criteria | If working with ADAS or AD product development, or if they had participated in a similar study before. |
| | Age | Between 24 - 68 years (M = 40.3, SD = 12.4). |
| | Gender | 12 females and 42 males. |
| | Experience with assistance | 13 participants had experience with Pilot Assist or similar level 2 systems.<br>23 participants had some experience with ACC and LKA.<br>18 no experience with these systems. |

Previous analysis in Streubel et al (2024) did not reveal any differences between participants with respect to age, gender, system experience dependent on whether they had a crash/near-crash or not.

# Drivers have impaired short-term memory under high workload: Safety Implications for transitions of control from vehicle automation

Rafael C. Gonçalves [1*], Jonny Kuo [2], Mike G. Lenné [2], Natasha Merat[1]

[1] University of Leeds | Institute for Transport Studies, Faculty of Environment, University of Leeds, Leeds LS2 9JT, United Kingdom
[2] Seeing Machines, VIC 3066, Melbourne, Australia
[*]trarg@leeds.ac.uk

**Abstract: The objective of this study was to evaluate the effect of both cognitive load and visual distraction on drivers' takeover performance during an SAE Level 2 automated drive. A driving simulator study was conducted where drivers needed to take over control during a safety critical scenario, while i) engaged in a 2-back task (cognitive load), ii) during ambient occlusion of the scene (visual distraction) or iii) a combination of both. Results show that drivers have their performance compromised by a combination of cognitive load and visual distraction. The results are useful for development of future driver state monitoring technologies.**

## 1. Introduction

### *1.1 Visual distraction and takeover performance*

Vehicle automation technology makes drivers more prone to visual distractions and non-driving related activities (Carsten et al., 2012), removing them from the control loop required for safe resumption of manual driving (Merat et al., 2019). However, regulations for level 2 (L2) vehicle automation (SAE, 2021) still require drivers to be able to resume control, whenever the automation reaches a system limitation.

Visual distraction has detrimental effects on drivers' safety (Young et al., 2009), also affecting their ability to retain the correct information about their surrounding environment (see Polani, 2011; Klyubin et al., 2007), reducing their situation awareness (SA, Endsley, 1995). Good SA during automation is also required for the safe resumption of control in case of system limitations (Merat et al., 2019). Results from L2 driving simulator studies have shown the negative impact of visual distraction on takeover performance (e.g., Zeeb et al., 2015, Louw et al., 2016, Li et al., 2021). The tendency for drivers to look away from the forward roadway during automated driving has been one motivation for the recommended integration of Driver Monitoring Systems (DMS) in vehicles (Euro NCAP, 2022). Based on these recommendations, camera-based DMS typically use the amount of time drivers take their eyes away from the roadway to warn drivers, or disengage the automated system.

### *1.2 Cognitive load and takeover performance*

Another factor that seems to be detrimental for drivers' takeover performance is their cognitive load. For manual driving, Engstrom et al. (2017) suggest that drivers' capability to detect context-relevant information about the road environment may be diminished by cognitive load. Liang & Lee (2010) and Broadbent et al. (2023) have suggested that drivers under cognitive load are less likely to scan the environment.

Yang et al. (2022) have reported that in L2 automation, drivers fail to detect peripheral visual stimuli that are more 20° away from the road centre. Drivers' gaze dispersion is also compromised by a cognitively-loading task (Wilkie et al., 2019). High cognitive load also affects takeover reaction time and vehicle control after a takeover (Du et al., 2020; Lu et al., 2021; Melnicuk et al., 2021).

### *1.3 Study aim*

The use of intermittent ambient occlusion is a popular technique for inducing visual distraction in driving simulator studies (Pettitt et al., 2006; Kujala et al., 2023). The technique is used to mimic the effects of drivers' long glances away from the road, when they engage with visual tasks, such as looking towards in-vehicle devices.

The n-back task has been used successfully to study the cognitive (non-visual) effect of in-vehicle tasks, such as hands-free mobile phone conversations (Mehler et al., 2011; Stojmenova & Sodnik, 2018).

Therefore, to study the effect of visual and cognitive distraction, and their combined influence on transition of control from L2 automation, drivers' take-over performance after a period of i) ambient occlusion, ii) 2-back task, and iii) the combination of the two tasks was investigated.

## 2. Method

### *2.1 Participants*

A total of 31 (13 female) drivers, aged from 22-56 years old (M = 38.02 years, SD = 12.03 years) took part in this study. All drivers had at least 3 years' experience driving in the UK, drove at least twice a week (average annual mileage of 12432 miles) and had no previous experience with vehicle automation.

### *2.2 Apparatus*

The University of Leeds Driving simulator, a 6-degree of freedom motion-based driving simulator, with a projection angle of 300 ° was used for this study.

### *2.3 Design and procedure*

The experiment followed a 3x2 repeated measures design, with the distraction manipulation (occlusion, 2-back, 2-back+occlusion), and event criticality (critical takeover, non-critical takeover) as the independent variables. Participants completed a single experimental drive, with a total of ten driving automation events, in a fully counterbalanced order.

Participants drove on a 3-lane motorway, with ambient surrounding traffic, assisted by an L2 automated system (SAE, 2021). The automation controlled both the lateral and longitudinal movement of the vehicle, keeping it in the middle of the centre lane, at a constant speed of 70 mph. Each of the ten events lasted 2 minutes and 20 seconds, and drivers were instructed to always monitor the environment. They were required to resume manual control at the end of each section, in response to a takeover request (TOR). Drivers were also told that they could resume control at any time if they felt the need to do so, for example, to avoid a potentially safety-critical situation. An auditory tone was used as the TOR, accompanied by a flashing red steering wheel icon, presented on the instrument cluster.

During each event, participants encountered one of the three types of distraction. Each distraction condition began as soon as automation was engaged, until the TOR was provided. Three of the ten automation events ended with a safety-critical situation (one for each of the three distraction conditions). The other seven were treated as ghost trials, to avoid a well-learnt response by drivers.

The safety critical situations were characterized by a hard brake from the lead vehicle in the middle lane, which the automated system was unable to manage. This lack of response from the automation created the likelihood of a collision at a time-to-collision (TTC) of 3s. For these events, the TOR was issued when the TTC was at 2s. The distraction conditions are outlined further below.

### *2.3.1 Ambient occlusion*

The occlusion manipulation (Senders et al., 1967) occurred every 9 seconds. For each 9-second period, the driving scene was overlayed with an opaque screen for 3s, as shown in **Fig 1**. The mirrors and the dash area, which included the instrument cluster, were all occluded during this 3-second period.

***Fig 1****: Example of the ambient occlusion.*

### *2.3.2 The 2-Back task*

An auditory version of the 2-back task (Mehler et al., 2011) was used to simulate a cognitively-loading condition, with no visual or manual element. Participants heard a series of random numbers, ranging from 0 to 9 through the car's speakers, presented every 2s over the course of the automated drive. Drivers were asked to repeat verbally, the second-to-last number they heard in the list.

## 3. Results

Drivers' takeover response to the critical events was measured using the time between the moment the lead vehicle started braking, to when drivers turned their steering wheel more than 2° or pressed the brake pedal over 1° (see Louw et al., 2018). To study the effect of each distraction manipulation on takeover performance, a one-way ANOVA was conducted on drivers' reaction time to the critical events. Results showed a significant effect of distraction manipulation [$F\ (4, 104) = 3.475$, $p = .019$, $\eta p2 = .197$], where drivers in the "2-back + occlusion" condition presented significantly higher reaction times compared to the other two conditions (**Fig 2**). The test also found a significant effect of event order as a covariant [$F\ (4, 104) = 16.354$, $p < .001$, $\eta p2 = .192$], where drivers were slower to react in their first trial, suggesting learning effects for the takeover response.

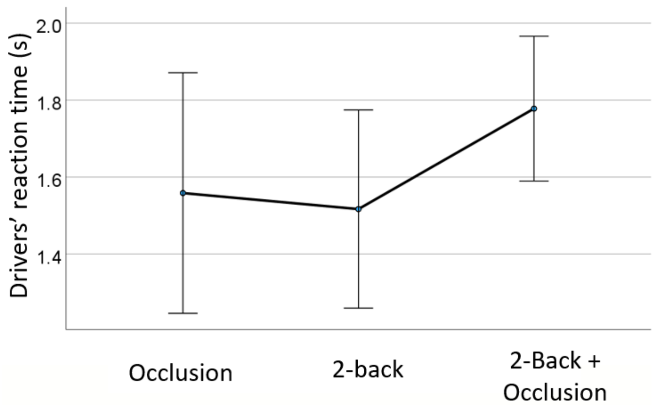



***Fig 2:*** *ANOVA results.*

To confirm whether drivers were successfully monitoring their environment and able to detect safety-critical events, we compared the likelihood of drivers reacting to the braking lead vehicle before the TOR was issued, across the 3 conditions. Chi-squared tests (**Fig 3**) showed a significant difference across the conditions [$X^2\ (2, 93) = 14.63$, $p=.001$], where no driver was able to react before the TOR for the "2-back + occlusion" condition.

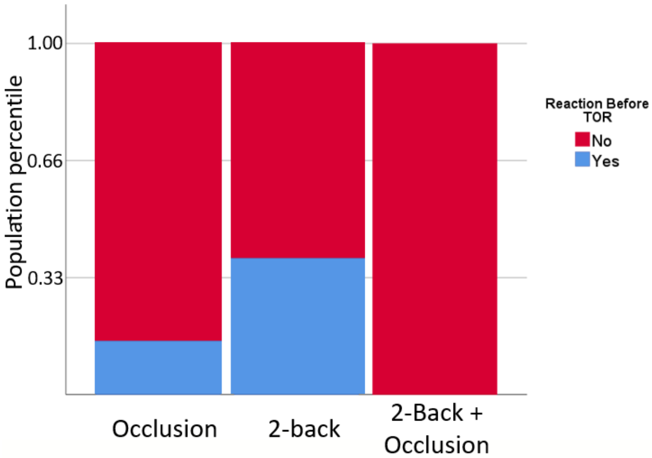



***Fig 3:*** *Chi-squared results.*

## 4. Discussion and Conclusions

This reaction time analysis of takeover performance suggests that the combination of both visual distraction and cognitive load together compromises drivers' reaction time to a TOR. Analysis of drivers' reaction to the lead vehicle before the TOR also suggests that the combined tasks were particularly challenging for drivers, suggesting that this condition reduced drivers' ability to respond to safety critical events.

Polany (2011) and Klyubin et al. (2007) have demonstrated that the taxation of drivers' working memory may affect their information processing capabilities of dynamic environments. The results of this study suggest that this diminished capability to process dynamic information magnifies the impairments in object detection, caused by cognitive load, as suggested by Engstrom et al.'s (2017) framework.

Capturing visual attention and driver engagement in consumer ratings represents a logical first step. The findings from this study suggest that cognitive load be considered for inclusion in future policy updates.

## 5. Acknowledgments

This research was funded by the Seeing Machines bi-lateral collaboration fellowship with the University of Leeds. The authors would like to thank Mr. Michael Daly, Mr. Anthony Horrobin, and Dr Albert Solernou, for developing the driving simulator experiment, and operating the driving simulator.

## References

Broadbent, D. P., D'Innocenzo, G., Ellmers, T. J., Parsler, J., Szameitat, A. J., & Bishop, D. T. (2023). Cognitive load, working memory capacity and driving performance: A preliminary fNIRS and eye tracking study. Transportation Research Part F: Traffic Psychology and Behaviour, 92, 121–132. https://doi.org/10.1016/J.TRF.2022.11.013

Carsten, O., Lai, F. C. H., Barnard, Y., Jamson, A. H., & Merat, N. (2012). Control Task Substitution in Semiautomated Driving. Human Factors: The Journal of the Human Factors and Ergonomics Society, 54(5), 747–761. https://doi.org/10.1177/0018720812460246

Du, N., Kim, J., Zhou, F., Pulver, E., Tilbury, D. M., Robert, L. P., Pradhan, A. K., & Yang, X. J. (2020). Evaluating Effects of Cognitive Load, Takeover Request Lead Time, and Traffic Density on Drivers' Takeover Performance in Conditionally Automated Driving. Proceedings - 12th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2020, 66–73. https://doi.org/10.1145/3409120.3410666

Engström, J., Markkula, G., Victor, T., & Merat, N. (2017). Effects of Cognitive Load on Driving Performance: The Cognitive Control Hypothesis. Human Factors, 59(5), 734–764. https://doi.org/10.1177/0018720817690639

Euro NCAP. (2022). Assisted Driving Highway Assist Systems.

Klyubin, A. S., Polani, D., & Nehaniv, C. L. (2007). Representations of space and time in the maximization of information flow in the perception-action loop. Neural Computation, 19(9), 2387–2432. https://doi.org/10.1162/NECO.2007.19.9.2387

Kujala, T., Kircher, K., & Ahlström, C. (2023). A Review of Occlusion as a Tool to Assess Attentional Demand in Driving. Human Factors, 65(5), 792–808. https://doi.org/10.1177/00187208211010953/ASSET/IMAGES/LARGE/10.1177_00187208211010953-FIG2.JPEG

Li, Q., Hou, L., Wang, Z., Wang, W., Zeng, C., Yuan, Q., & Cheng, B. (2021). Drivers' visual-distracted take-over performance model and its application on adaptive adjustment of time budget. Accident Analysis & Prevention, 154, 106099. https://doi.org/10.1016/J.AAP.2021.106099

Liang, Y., & Lee, J. D. (2010). Combining cognitive and visual distraction: Less than the sum of its parts. Accident Analysis & Prevention, 42(3), 881–890. https://doi.org/10.1016/J.AAP.2009.05.001

Louw, T., Madigan, R., Carsten, O., & Merat, N. (2016). Were they in the loop during automated driving? Links between visual attention and crash potential. Injury Prevention : Journal of the International Society for Child and Adolescent Injury Prevention, 23(4), 281–286. https://doi.org/10.1136/injuryprev-2016-042155

Louw, T., Markkula, G., Boer, E. R., Madigan, R., Carsten, O. M. J., & Merat. (2018). Coming back into the loop: Drivers' perceptual-motor performance in critical events after automated driving. Accident Analysis & Prevention, 108, 9–18. https://doi.org/10.1016/J.AAP.2017.08.011

Lu, G., Zhai, J., Li, P., Chen, F., & Liang, L. (2021). Measuring drivers' takeover performance in varying levels of automation: Considering the influence of cognitive secondary task. Transportation Research Part F: Traffic Psychology and Behaviour, 82, 96–110. https://doi.org/10.1016/J.TRF.2021.08.005

Mehler, B., Reimer, B., & Dusek, J. A. (2011). MIT AgeLab Delayed Digit Recall Task (n-back)(Paper 2011-3B)(2011-06-28). https://www.researchgate.net/publication/230729111

Melnicuk, V., Thompson, S., Jennings, P., & Birrell, S. (2021). Effect of cognitive load on drivers' State and task performance during automated driving: Introducing a novel method for determining stabilisation time following take-over of control. Accident Analysis & Prevention, 151, 105967. https://doi.org/10.1016/J.AAP.2020.105967

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), 87–98. https://doi.org/10.1007/s10111-018-0525-8

Pettitt, M. A., Burnett, G. E., Bayer, S., & Stevens, A. (2006). Assessment of the occlusion technique as a means for evaluating the distraction potential of driver support systems. IEE Proceedings: Intelligent Transport Systems, 153(4). https://doi.org/10.1049/IP-ITS:20060027

Polani, D. (2011). An informational perspective on how the embodiment can relieve cognitive burden. IEEE SSCI 2011 - Symposium Series on Computational Intelligence - IEEE ALIFE 2011: 2011 IEEE Symposium on Artificial Life, 78–85. https://doi.org/10.1109/ALIFE.2011.5954666

SAE. (2021). Summary Of SAE International'S Levels Of Driving Automation For On-Road Vehicles. https://www.sae.org/misc/pdfs/automated_driving.pdf

Senders, J. W., Kristofferson, A. B., University, M., Hamilton, O. ;, Levison, W. H., Dietrich, C. W., Ward, J. L., & Inc, N. (1967). The Attentional Demand of Automobile Driving.

Stojmenova, K., & Sodnik, J. (2018). Detection-Response Task—Uses and Limitations. Sensors 2018, Vol. 18, Page 594, 18(2), 594. https://doi.org/10.3390/S18020594

Wilkie, R., Mole, C., Giles, O., Merat, N., Romano, R., & Markkula, G. (2019). Cognitive Load During Automation Affects Gaze Behaviours and Transitions to Manual Steering Control. Driving Assessment Conference, 10(2019), 426–432. https://doi.org/10.17077/DRIVINGASSESSMENT.1728

Yang, S., Wilson, K., Roady, T., Kuo, J., & Lenné, M. G. (2022). Beyond gaze fixation: Modeling peripheral vision in relation to speed, Tesla Autopilot, cognitive load, and age in highway driving. Accident Analysis & Prevention, 171, 106670. https://doi.org/10.1016/J.AAP.2022.106670

Young, K. L., Regan, M. A., & Lee, J. D. (2009). Measuring the Effects of Driver Distraction: Direct Driving Performance Methods and Measures.

Zeeb, K., Buchner, A., & Schrauf, M. (2015). What determines the take-over time? An integrated model approach of driver take-over after automated driving. Accident Analysis & Prevention, 78, 212–221. https://doi.org/10.1016/j.aap.2015.02.023

**Investigating Changes in Cognitive Workload and Glance Allocation during Partially Automated Driving in Construction Zones**

Noor Jajo [1*], & Francesco Biondi, Ph.D. [1]

[1] Human Kinetics, University of Windsor, 2555 College Ave, Windsor, Canada
[*]jajon@uwindsor.ca

**Partially automated vehicles pose concerns for driver disengagement and distraction, especially in construction zones, prompting this study's investigation into their impact on visual attention. Using a factorial design, the study examines driving modes (manual and level 2) and zones (preconstruction, construction, post-construction) with dependent measures including total glance time per Area of interest (AOI) and total eyes-off-the-road time (TEORT). Thirty participants drove a Tesla Model 3 on Ontario Highway 401, with visual attention examined through video coding. Results show that during partially automated driving, participants spent more time glancing at the touchscreen, especially in construction zones. However, time spent checking mirrors remained consistent across modes and zones. These findings suggest potential distractions and increased non-driving related tasks during partially automated driving, highlighting the need for further research in this area.**

## 1. Introduction

The levels of automated driving systems, as defined by the Society of Automotive Engineers, range from level 0 (manual) to level 5 (full automation) (SAE, 2021). Level 2 or partially automated systems are responsible for tasks such as accelerating, braking, and steering, with the human driver being in charge of supervising the system and taking control when necessary.

With the role of the driver transitioning from vehicle operator to system supervisor, a Human Factors concern with partially automated systems is that drivers may be more likely to disengage from the supervisory task when the automated system is engaged. Recent studies show that, when in partially automated mode, drivers may experience a reduction in cognitive workload possibly as a result of increased boredom (Biondi et al., 2023; McWilliams & Ward, 2021). It follows that, instead of continuing to pay attention to the road, they may be more inclined to engage in potentially distracting activities (Noble et al., 2021), which are known to have a negative impact on road safety.

The issue of to what extent partially automated driving is detrimental to cognitive workload and attention allocation is even more important when operating these vehicles in construction zones. Construction zones present unique challenges, including reduced road visibility, unexpected traffic changes, and speed fluctuations. Recent data from the National Safety Council (2023) show that fatalities in construction zones have risen by 60% since 2010. It follows that for partially automated systems to be deemed safe, more needs to be known on how their use influences drivers' workload and attention allocation in construction zones.

This study tackles this issue by measuring drivers' behavior when at the wheel of a partially automated vehicle through a construction zone. Participants drive a 2022 Tesla Model 3 on a section of Ontario Highway 401 in either manual or partially automated mode through construction and non-construction zones. Visual attention is assessed via measuring drivers' glance time and total eyes-off-the-road-time (TEORT).

## 2. Method

### 1.1 *Participants*

Thirty participants (13 females, average age 22 ± 4.36) were recruited from the University of Windsor. Eligible participants held valid driver's licenses, had no at-fault accidents in the past 2 years, had corrected-to-normal vision, completed a 45-minute defensive driving course, and were fluent in English.

### 1.2 *Design*

A factorial design with 2 independent variables were used in this study: driving modes (2 levels: manual and L2), and zones (3 levels: preconstruction, construction, post construction). The dependent measures are total glance time by area of interest (AOI), total eyes off road time (TEORT).

### 1.3 *Equipment*

A Tesla Model 3 equipped with adaptive cruise control, lane-keep assist, automated steering, and automated acceleration and braking was used. Three GoPro HERO8 Black cameras recorded participant view, front view capturing construction zones, and touchscreen view (See Figure 1).


*Figure 1: (A) Participant View, (B) Front View, (C), Touchscreen View.*

*1.4 Route*

Participants familiarized themselves with the vehicle by driving around the University of Windsor campus before the experiment. They then drove on a section of Ontario Highway 401 in both manual and L2 mode (counterbalanced across participants and each lasted for 40 minutes). The research associate was always seated in the back.

*1.5 Glances*

Both participant and front view videos were manually coded. Two researchers coded participant view videos to identify four AOIs: front road, touchscreen, side mirrors, and rearview mirror. See Figure 2 for more details. Discrepancies were resolved by a third coder. An inter-rater reliability analysis yielded a strong Cohen's Kappa of 0.85 between the two coders. The front view videos were coded by a single researcher to identify construction zones, including preconstruction 1, construction 1, postconstruction 1, construction 2, postconstruction 2, construction 3, postconstruction 3, and construction sign. However, data analysis excluded zones except preconstruction 1, construction 1, and postconstruction 1, due to limited participant experiences.

| **Glance Coding Definitions** | |
|---|---|
| **Front road (Blue)** | Any glance made at the forward roadway even if they look at the center or right to inspect the road |
| **Touchscreen (Purple)** | Any glance made at the instrument panel |
| **Side mirrors (green)** | Any glance made the left and right-side mirrors |
| **Rearview mirrors (Yellow)** | Any glance made at the rearview mirror. |

*Figure 2: AOI distribution and definitions.*

*1.6 Procedure*

Participants were provided with a Tesla Model 3 Autopilot training video three days prior to the study to familiarize themselves with the system. Upon arriving at the garage where the Tesla was parked, participants were screened for alcohol, caffeine, and drug use. Those with recent consumption or excessive intake were excluded. They then received a detailed study introduction, including procedures, risks, and benefits. Following that Participants were familiarized with the vehicle by adjusting mirrors, seat, and steering wheel, and receiving instructions for manual and L2 driving.

*1.7 Statistical Analysis*

Bayes Factor analysis was used to determine the effect of the independent variables on the outcome variables. Bayes factor analysis, chosen over traditional NHST, provides direct evidence against the null hypothesis. It offers three conclusions (support for null, support for alternative, and weak evidence without clear support) and quantifies the strength of evidence. The Bayes Factor (BF) determines the likelihood of the data under either hypothesis, with BF greater than 10 or less than 0.1 indicating strong evidence for the alternative or null hypothesis, respectively (Jeffrey 1935). Between 0.1 and 10 shows weak evidence for either hypothesis. Unlike NHST, BF analysis offers a more nuanced assessment of evidence. RStudio with BayesFactor package was used for this analysis.

**3. Results**

The aim of this analysis was to explore changes in visual attention during partial automation across various construction zones. Figure 2 illustrates the average percentage of time spent in each AOI and construction zone under different driving modes. See Appendix A for descriptive statistics.

A Bayesian ANOVA was used to analyze the effect of mode and construction zone on average percent in each AOI. Results revealed significant effects of driving mode on TEORT, supported by extreme evidence (BF = 5462.32), with very strong evidence for the interaction between driving mode and construction zone (BF = 96.13). Although the influence of construction zone alone on TEORT was not significant, moderate evidence supported its effect (BF = 0.08), along with extreme evidence for the combined effect of driving mode and construction zone (BF = 477.05). Similar patterns were observed for the touchscreen, where driving mode significantly influenced time spent looking at it (BF = 4206.01), with very strong evidence for the interaction with both construction zone and driving mode (BF = 58.78). Construction zones alone showed moderate evidence for their effect (BF = 0.08). However, the impact on side and rearview mirrors differed, with weak evidence for driving mode's effect on side mirrors (BF = 1.82) and no effect on rearview mirrors (BF = 0.65). Similarly, construction zones had no significant effect on mirrors, supported by moderate evidence (BF = 0.1 for side mirrors, BF = 0.08 for rearview mirrors), as did their combined effect with driving mode (BF = 0.20 for side mirrors, BF = 0.05 for rearview mirrors).


*Figure 3: Bar plot of average % of time spent in each AOI by construction zone in both driving modes.*

## 4. Discussion

This study aimed to assess the impact of partially automated driving and construction zones on driver visual attention allocation. Results indicate a notable increase in TEORT during partially automated driving, particularly in construction zones, with participants averting their gaze 9.12% of the time compared to 3.2% during manual driving. While time spent checking side and rearview mirrors remained consistent across driving modes and zones, there was a notable rise in glances made at the touchscreen. Specifically, during L2 driving in construction zones, participants spent an average of 6.44% of the time glancing at the touchscreen, compared to 2.17% during manual driving. These findings are consistent with prior research by Zangi et al. (2022), suggesting increased engagement in non-driving related tasks (NDRT) during partially automated driving. Their study demonstrated that NDRT engagement during partially automated driving significantly impairs hazard perception, leading to fewer hazard identifications and glances. While this corresponds with the expectation of drivers potentially becoming complacent when driving with partially automated mode, it is also plausible that participants are spending more time looking at their touchscreen during partially automated and construction due to monitoring the system. The system did alert participants that partially automated mode would disengage since the lane markings were unclear. However, after the construction zone was complete their attention did not go back to the front road in the partially automated mode, they spent 8.94% of their total time glancing away from the road as opposed to 4.21% in manual mode. Results also showed that during all AOIs none were affected by construction zone alone, but the evidence is moderate.

While these findings add to the understanding of drivers' behaviour changes with partially automated driving, it is important to note some limitations. Our study only looked at one construction zone out of the three identified due to not all participants experiencing all construction zones. Therefore, this data does not represent all the 2 hours driven, but only a portion of it. Future research should consider a controlled study that accounts for all participants experiencing different construction zones. In addition, this study only looked at investigating visual attention. Research has shown a link between visual attention and cognitive workload, with participants spending more time engaging in non-driving tasks due to boredom (Strayer & Fisher, 2016).

## 5. Conclusions

This study is among the first to examine how partially automated driving in construction zones affects visual attention. Our results revealed a rise in touchscreen glances and TEORT during partially automated mode in construction zones. While this study only covers a portion of the drive, future research could explore the full impact of partially automated on visual attention in construction settings.


## 6. Acknowledgments

We thank the Ontario Ministry of Transportation, SSHRC, and NSERC for their funding.


## References


Biondi, F. N., McDonnell, A. S., Mahmoodzadeh, M., Jajo, N., Balasingam, B., & Strayer, D. L. (2023). Vigilance Decrement During On-Road Partially Automated Driving Across Four Systems. *Human Factors*. https://doi.org/10.1177/00187208231189658

Jeffreys H. (1935). Some Tests of Significance, Treated by the Theory of Probability. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(2), 203–222.

McWilliams, T., & Ward, N. (2021). Underload on the Road: Measuring Vigilance Decrements During Partially Automated Driving. *Frontiers in Psychology*, *12*(April), 1–13. https://doi.org/10.3389/fpsyg.2021.631364

National Safety Council. (2023). Motor Vehicle Safety Issues Work Zones. Injury Facts NSC. Retrieved March 2, 2024, from https://injuryfacts.nsc.org/motor-vehicle/motor-vehicle-safety-issues/work-zones

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 and Prevention*, *151*(March 2020), 105959. https://doi.org/10.1016/j.aap.2020.105959

SAE. (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. https://www.sae.org/standards/content/j3016_202104/

Strayer, D. L., & Fisher, D. L. (2016). SPIDER: A Framework for Understanding Driver Distraction. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 58(*1), 5–12. https://doi.org/10.1177/0018720815619074

Zangi, N., Srour-Zreik, R., Ridel, D., Chassidim, H., & Borowsky, A. (2022). Driver distraction and its effects on partially automated driving performance: A driving simulator study among young-experienced drivers. *Accident Analysis & Prevention*, *166*, 106565. https://doi.org/10.1016/j.aap.2022.106565

## Appendix A

Descriptive statistics of the average percent of time spent in each AOI and construction zone by mode. Mean and standard deviation (SD) are included.

| L2 Mode | | | |
|---|---|---|---|
| **AOI** | **Construction Zone** | **Mean (%)** | **SD** |
| **Front Road** | C1 | 90.8817638 | 8.443452 |
| **Front Road** | PostC1 | 91.0614588 | 7.7769709 |
| Front Road | PreC1 | 92.6964054 | 4.6728529 |
| Rearview Mirrors | C1 | 0.6196172 | 1.0750927 |
| Rearview Mirrors | PostC1 | 0.8622677 | 0.9218863 |
| Rearview Mirrors | PreC1 | 0.8492432 | 1.1559533 |
| Side mirrors | C1 | 1.583456 | 4.5059044 |
| Side mirrors | PostC1 | 1.5014659 | 1.9120269 |
| Side mirrors | PreC1 | 0.797333 | 1.0131959 |
| Touchscreen | C1 | 6.4431926 | 5.5440541 |
| Touchscreen | PostC1 | 6.4420201 | 5.9678039 |
| Touchscreen | PreC1 | 5.6176139 | 3.9622046 |
| **Manual Mode** | | | |
| **AOI** | **Construction Zone** | **Mean (%)** | **SD** |
| **Front Road** | C1 | 96.8072206 | 3.1085401 |
| **Front Road** | PostC1 | 95.7994619 | 4.042855 |
| Front Road | PreC1 | 95.8860299 | 4.1424215 |
| Rearview Mirrors | C1 | 0.546681 | 0.7786921 |
| Rearview Mirrors | PostC1 | 0.4855777 | 0.653759 |
| Rearview Mirrors | PreC1 | 0.5430793 | 0.711732 |
| Side mirrors | C1 | 0.3938245 | 0.5456739 |
| Side mirrors | PostC1 | 0.5811187 | 1.1874089 |
| Side mirrors | PreC1 | 0.4609881 | 0.6715558 |
| Touchscreen | C1 | 2.1728321 | 2.3695262 |
| Touchscreen | PostC1 | 3.0986047 | 3.2752481 |
| Touchscreen | PreC1 | 3.093355 | 3.251724 |

# Advanced Distracted Driving Detection: Deep Learning Methods for Cell Phone Usage Recognition

Zifei Wang [1], Meitang Li [2], Huizhong Guo [2], Carol Flannagan [2], Jim Sayer[2], Shan Bao[1,2*]

[1] Department of Industrial and Manufacturing Systems Engineering, University of Michigan-Dearborn, 4901 Evergreen Rd, Dearborn, MI, USA, 48128
[2] University of Michigan Transportation Research Institute, 2901 Baxter Rd, Ann Arbor, MI, USA, 48109
[*]shanbao@umich.edu

**Abstract:**
**Distracted driving is a critical threat to road safety, leading to numerous motor vehicle crashes and fatalities. With the widespread of in-vehicle infotainment systems, particularly cell phones, the issue of driver distraction is increasingly exacerbated. To address the issue, this research aims to enhance the detection of distracted driving, specifically focusing on the usage of cell phones, through an integration of computer vision and deep learning methods. By leveraging the IVBSS dataset, which contains both driving videos and vehicle kinematic features, we aim to improve the effectiveness and robustness of driver distraction detection. The study employed OpenFace to identify facial features of drivers and YOLOv4 for object detection on cell phones. Three deep learning methods, including Hidden Markov Model, Support Vector Classification, and Long Short-Term Memory, are implemented to kinematic features extracted from vehicle data for early distraction classification. Additionally, the deep learning methods are compared and selected based on prediction performance and further modified to be trained with an extra input of the prediction result from computer vision methods. The fusion of these techniques can improve the detection of cell phone usage during driving scenarios, with promising accuracy and F1 scores. The findings suggest that integrating computer vision techniques with deep learning models can yield a robust framework for detecting distracted driving behaviors.**

## 1. Introduction

Driver distraction is a serious public safety concern. In 2021, 3522 people were killed in distraction-involved crashes (NCSA, 2023). The rise of in-vehicle infotainment systems, like cell phones, has introduced new distractions, increasing crash potential. Texting, for example, diverts drivers' eyes, hands, and attention from driving (Simons-Morton et al., 2014).

Numerous distraction detection systems have been developed to monitor driver's behavior and mitigate potential crashes. Common approaches included image-based classification of head and eye gaze using interior cameras (Miyaji et al., 2009; Liu et al., 2015; Hari & Sankaran, 2021). Posture and hand position are also key research areas for detecting manual and visual distraction (Chang et al., 2020; Li et al., 2020). Driver's cognitive state could also significantly influence driving performance (Kanaan et al., 2019). Abnormal mean shifts of lateral speeds, steering wheel reversal rate, and prediction errors of lane positions were indicators of distracted driving (Li et al., 2017; Kountouriotis et al., 2016).

The objective of this study is to identify whether drivers are engaged in secondary tasks (e.g., cell phone use) through computer vision and deep learning methods. Both video image data vehicle kinematic features are utilized to predict driver distraction. The best-performing deep learning methods were further modified with an extra input, the prediction result of the computer vision results.

## 2. Method

### *2.1 Data*

This study used data from UMTRI's Integrated In-vehicle Based Safety System Study (IVBSS) (Sayer et al., 2011). We analyzed 932 manually labelled distracted driving events from 58 drivers across 300 video clips. Valid clips were manually coded into binary variables indicating phone usage (true or false).

We extracted face videos (head and eye positions), cabin videos (hand movements), and vehicle kinematic features (longitudinal speed, longitudinal acceleration, steering angle, and lateral speed). The data were saved at a 10 Hz resolution, with 0.1-second intervals to capture driving dynamics accurately. Feature extraction was performed automatically using IVBSS tools and scripts.

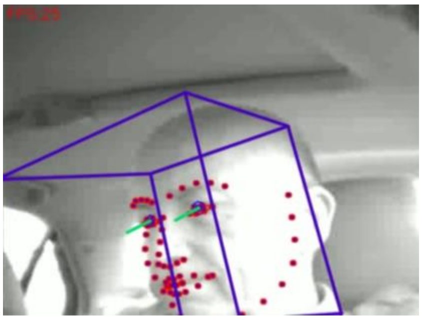
### *2.2 Video Image Processing*

#### *2.2.1 OpenFace:*

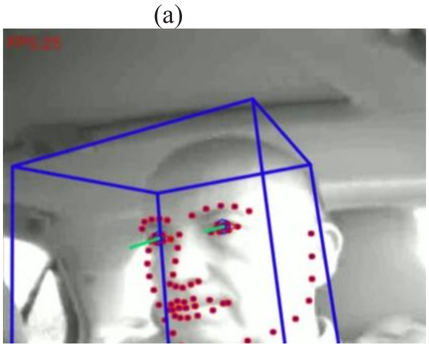
OpenFace is a neural network-based face recognition tool in Python (Schroff et al., 2015). In the analysis, OpenFace was employed to identify driver face positions and calculate corresponding head pose locations. Furthermore, we used coordinate transformation to identify the relative eye-gaze locations regarding driver's head pose (see Fig.1).

#### *2.2.1 YOLOv4:*

A limitation of OpenFace was that it relied on gaze directions, which ignored distraction when drivers' face looking straight (e.g., phone conversation). The YOLOv4 image processing tool was explored to detect phone locations near the driver's body (Bochkovskiy et al., 2020). An example of phone detection from the cabin videos are shown below (see Fig. 2).



(a)



(b)

***Fig. 1.*** *Examples of driving episodes detected by OpenFace* ***(a)*** *distracted driving episode (looking off the roads),* ***(b)*** *Non-distracted driving (looking on the roads))*


***Fig. 2.*** *Detected cell phone use of a driver by YOLOv4*

### *2.3 Kinematic Feature Prediction*

Three deep learning methods were applied on the vehicle kinematic features, Hidden Markov Model (HMM), Support Vector Classification (SVC), and Long Short-Term Memory (LSTM). Each driver was assigned a model specifically trained on their data.

#### *2.3.1 Hidden Markov Model:*

The HMM is statistics-based model to inference hidden states from influenced observations (see Fig. 3). The hidden state was a discrete variable indicating distraction state.

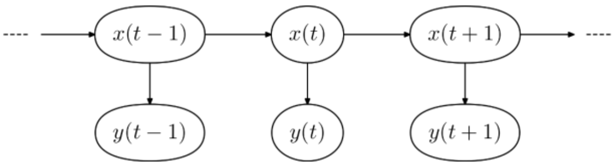



***Fig. 3.*** *Hidden Markov Model*

#### *2.3.2 Support Vector Classification:*

SVC is a machine learning algorithm widely used for classification (See Fig. 4). Each data sample contains 5, 10 or 15 timestamps. The distraction state is 5 timestamps after the last one of the input data for early detections.

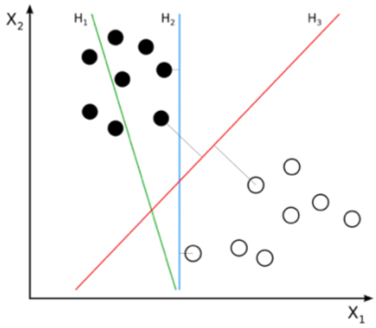



***Fig. 4.*** *Support Vector Classification Description*

#### *2.3.3 Long Short-Term Memory:*

LSTM is a recurrent neural network architecture to model sequences and long-term dependencies effectively. LSTM cells include three main gates: forget ($f_t$), input ($e_t$), and output ($o_t$) gates, which regulate the flow of information and help preserve the error that can be backpropagated through time and layers (see Fig. 5). This architecture helps to mitigate the vanishing error problem, allowing the model to learn long-term dependencies in sequential data (Staudemeyer et al., 2019).

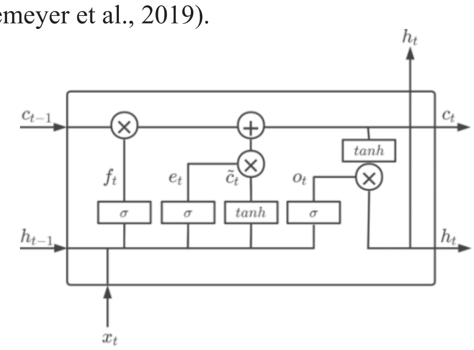



***Fig. 5.*** *LSTM Model Structure (Fu et al., 2022)*

## 3. Results

### *3.1 Video Processing*

The performance of OpenFace and YOLOv4 were shown in Table 1. In the combined results, if either model was detected "distracted" for the data sample, the final output will be "distracted."

***Table 1*** *Video Processing Results*

| | Accuracy | F1 | Precision | Recall |
|---|---|---|---|---|
| OpenFace | 0.621 | 0.666 | 0.711 | 0.725 |
| YOLOv4 | 0.344 | 0.122 | 0.449 | 0.086 |
| Combined | 0.628 | 0.676 | 0.710 | 0.739 |

### 3.2 Kinematic Feature Prediction

The four vehicle kinematic features were analyzed through deep learning models, with parameters tuned via 3-fold cross-validation. Table 2 shows the best models of two hidden states: {Distracted, Not Distracted} and three hidden states: {Pre-Distracted, Distracted, Post-Distracted}. Due to poor performance, the HMM was not considered further.

***Table 2*** *HMM Results*

| | Accuracy | F1 | Precision | Recall |
|---|---|---|---|---|
| 2 states | 0.431 | 0.370 | 0.580 | 0.348 |
| 3 states | 0.417 | 0.340 | 0.572 | 0.311 |

Table 3 showed SVC performance across three different setups. The 15-timestamps model achieved the best performance. However, there existed a 20% gap between precision, signaturing biased results.

***Table 3*** *SVC Results*

| | Accuracy | F1 | Precision | Recall |
|---|---|---|---|---|
| 5 timestamps | 0.751 | 0.788 | 0.720 | 0.931 |
| 10 timestamps | 0.761 | 0.788 | 0.733 | 0.916 |
| 15 timestamps | 0.784 | 0.806 | 0.745 | 0.935 |

Table 4 describes the LSTM model results. The longest timestamp achieved the best performance. While not promoting recall, LSTM increased the precision by 6%, showing its superiority in analyzing biased datasets.

***Table 4*** *LSTM Results*

| | Accuracy | F1 | Precision | Recall |
|---|---|---|---|---|
| 5 timestamps | 0.813 | 0.829 | 0.778 | 0.920 |
| 10 timestamps | 0.826 | 0.841 | 0.790 | 0.932 |
| 15 timestamps | 0.845 | 0,854 | 0.810 | 0.937 |

To improve accuracy, the combined video prediction results was included as an input (see Table 5). While the SVC performance did not improve much, the LSTM benefited greatly. The best model of LSTM with 15 timestamps and video results achieved 91% accuracy while the F1, precision and recall retained similar, showing that the bias was further decreased with the video prediction results.

***Table 5*** *Combination of Video Data and Driving Data Results*

| | Accuracy | F1 | Precision | Recall |
|---|---|---|---|---|
| SVC 5 | 0.751 | 0.788 | 0.720 | 0.931 |
| SVC 5+* | 0.754 | 0.791 | 0.723 | 0.933 |
| SVC 10 | 0.761 | 0.788 | 0.733 | 0.916 |
| SVC 10+ | 0.771 | 0.801 | 0.733 | 0.935 |
| SVC 15 | 0.784 | 0.806 | 0.745 | 0.935 |
| SVC 15+ | 0.788 | 0.810 | 0.746 | 0.937 |
| LSTM 5 | 0.813 | 0.829 | 0.778 | 0.920 |
| LSTM 5+ | 0.870 | 0.869 | 0.839 | 0.923 |
| LSTM 10 | 0.826 | 0.841 | 0.790 | 0.932 |
| LSTM 10+ | 0.887 | 0.883 | 0.857 | 0.931 |
| LSTM 15 | 0.845 | 0,854 | 0.810 | 0.937 |
| LSTM 15+ | 0.910 | 0.904 | 0.882 | 0.944 |

*: "+" indicates adding video prediction as a variable

Extra variables were added to the feature space for better predictions, including longitudinal speed, longitudinal acceleration, lateral speed, lateral acceleration, lane offset, yaw rate, the variance within 1 second for the variables above, traffic and video prediction. The comparison of the performance given the old and new set of variables were shown in Table 6. The LSTM benefited most with an 4.7% increment in accuracy.

***Table 6*** *Additional Driving Data Results*

| | Accuracy | F1 | Precision | Recall |
|---|---|---|---|---|
| SVC (10,old) | 0.771 | 0.801 | 0.733 | 0.935 |
| SVC (10,new) | 0.772 | 0.813 | 0.747 | 0.953 |
| SVC (15,old) | 0.788 | 0.810 | 0.746 | 0.937 |
| SVC (15,new) | 0.797 | 0.826 | 0.768 | 0.944 |
| LSTM (5,old) | 0.870 | 0.869 | 0.839 | 0.923 |
| LSTM (5,new) | 0.911 | 0.904 | 0.895 | 0.923 |
| LSTM (10,old) | 0.887 | 0.883 | 0.857 | 0.931 |
| LSTM (10,new) | 0.944 | 0.940 | 0.930 | 0.958 |
| LSTM (15,old) | 0.910 | 0.904 | 0.882 | 0.944 |
| LSTM (15,new) | 0.957 | 0.954 | 0.946 | 0.969 |

A permutation importance analysis was conducted using the best LSTM to evaluate feature contributions (see Fig. 6). The "baseline" represents the model's accuracy without permutation, serving as a reference point. Video prediction was the most important feature, followed by the variance of lateral acceleration and yaw rate. Features below "traffic" had similar importance, while the remaining features had comparatively smaller feature importance.

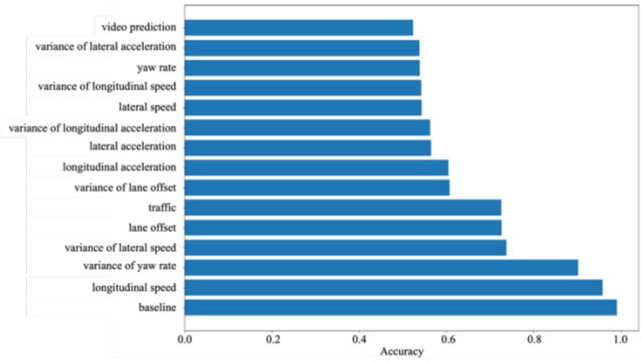



**Fig 6.** Permutation Feature Importance

## 4. Discussion

In general, the video only processing methods do not provide accurate prediction results. There are two possible causes and potential improvements:

- The IVBSS video dataset was monochromatic and of low resolution. It is expected that the algorithms will perform better on colored video with higher resolution.
- Manual coding of distracted video data can be improved, by considering the end of a distracting behavior for video labelling.

Apart from video prediction, deep learning methods on the non-video driving data have displayed varied performances.

- HMM, previously validated, achieved around 43% accuracy, possibly due to inconsistency of distraction event length.

- SVC models of 15 timestamps had highest accuracy, 78.4%. The recall is 93.5% but precision was only 74.5%, indicating a bias in learning.
- LSTM (15 timestamps) outperforms SVC with an average of 83% accuracy, improving precision to 81.0% while maintaining 93.7% recall, reducing bias.

The results of non-video driving data are better than the video only processing results, mostly because it is completely accessible to the deep learning model, while video data is limited by resolution and color. To further improve model results, video processing results were included as input variables in SVC and LSTM for comparisons:

- SVC performance did not significantly improve with the video prediction results.
- LSTM performance increased by 6% in accuracy with video predictions. The best model achieved 88.2% precision, and 94.4% recall. This means a smaller bias in the prediction.

## 5. Conclusions

In summary, this study analyzed cell phone-related distracted driving by integrating computer vision and deep learning methods. We have proposed a unique approach to combine video data with vehicle kinematic driving data from the IVBSS dataset, resulting in the best performance. Due to video resolution and color limitations, pure video processing yielded less accurate results, with OpenFace and YOLOv4 achieving around 63% accuracy. Deep learning methods on non-video data performed better, with SVC and LSTM models at 15 timestamps reaching 78.4% and 84.5% accuracy, respectively.

To enhance the prediction performance, the results of the video image processing were integrated as an input variable in the SVC and LSTM models. The LSTM accuracy reached 91% by leveraging the video data. After additional kinematics data extracted from the IVBSS dataset, the best-performing model's performance was pushed to 95.7%. The integration of the computer vision and deep learning methods significantly enhanced the robustness of cell phone-related distraction predictions. The findings of the research provide valuable insights on the enhancement of driving distraction detection for better in-vehicle detection system design, which could ultimately reduce the phenomenon of distraction and ensure a safer road environment for all.

## 6. Acknowledgments

The authors would like to acknowledge the US Department of Transportation for sponsoring this project.

## References

National Center for Statistics and Analysis. (2023, May). Distracted driving in 2021 (Research Note. Report No. DOT HS 813 443). National Highway Traffic Safety Administration.

Simons-Morton, B. G., Guo, F., Klauer, S. G., Ehsani, J. P., & Pradhan, A. K. (2014). Keep your eyes on the road: Young driver crash risk increases according to duration of distraction. Journal of Adolescent Health, 54(5), S61-S67.

Miyaji, M., Kawanaka, H., & Oguri, K. (2009, October). Driver's cognitive distraction detection using physiological features by the adaboost. In 2009 12th International IEEE conference on intelligent transportation systems (pp. 1-6). IEEE.

Liu, T., Yang, Y., Huang, G. B., Yeo, Y. K., & Lin, Z. (2015). Driver distraction detection using semi-supervised machine learning. IEEE transactions on intelligent transportation systems, 17(4), 1108-1120.

Hari, C. V., & Sankaran, P. (2021). Driver distraction analysis using face pose cues. Expert Systems with Applications, 179, 115036.

Chang, R. C. H., Wang, C. Y., & Shen, C. H. (2020, September). Modified YOLOv3-Tiny Using Dilated Convolution for Driver Distraction Detection. In 2020 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan) (pp. 1-2). IEEE.

Li, L., Zhong, B., Hutmacher Jr, C., Liang, Y., Horrey, W. J., & Xu, X. (2020). Detection of driver manual distraction via image-based hand and ear recognition. Accident Analysis & Prevention, 137, 105432.

Kanaan, D., Ayas, S., Donmez, B., Risteska, M., & Chakraborty, J. (2019). Using naturalistic vehicle-based data to predict distraction and environmental demand. International Journal of Mobile Human Computer Interaction (IJMHCI), 11(3), 59-70.

Li, Z., Bao, S., Kolmanovsky, I. V., & Yin, X. (2017). Visual-manual distraction detection using driving performance indicators with naturalistic driving data. IEEE Transactions on Intelligent Transportation Systems, 19(8), 2528-2535.

Kountouriotis, G. K., Spyridakos, P., Carsten, O. M., & Merat, N. (2016). Identifying cognitive distraction using steering wheel reversal rates. Accident Analysis & Prevention, 96, 39-45.

Sayer, J., LeBlanc, D., Bogard, S., Funkhouser, D., Bao, S., Buonarosa, M. L., & Blankespoor, A. (2011). Integrated vehicle-based safety systems field operational test: Final program report (No. FHWA-JPO-11-150; UMTRI-2010-36). United States. Joint Program Office for Intelligent Transportation Systems.

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 815-823).

Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

Fu, X., Meng, H., Wang, X., Yang, H., & Wang, J. (2022). A hybrid neural network for driving behavior risk prediction based on distracted driving behavior data. PLoS one, 17(1), e0263030.

Staudemeyer, R. C., & Morris, E. R. (2019). Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. arXiv preprint arXiv:1909.09586.

# Driver Distraction Detection using driver's focus of attention and eye movement

Yuto Kumamoto [1*], Kento Ohtani [1], Daiki Suzuki [2], Minori Yamataka [2], Kazuya Takeda [1]

[1] Graduate School of Informatics, Nagoya University, Aichi, Japan
[2] Advanced Research and Innovation Center, DENSO Corporation, Aichi, Japan
[*] kumamoto.yuto@g.sp.m.is.nagoya-u.ac.jp

**Abstract: Since most road accidents are caused by human error, researchers have tried to detect the factors causing these errors, and one of the factors that has been identified is distracted driving. Various methods of detecting and addressing this problem have been proposed in previous studies, however accurate driver distraction detection has proven to be difficult to achieve. In part, this is because there are several kinds of distractions, and because appropriate driving behavior can vary depending on the situation. Although distraction detection methods focusing on changes in the driver's eye movement, or the driver's focus of attention predicted from dashcam videos, have been proposed, these methods do not simultaneously consider both phenomena. Therefore, in this paper we proposed a distraction detector that uses both eye movement and direction of the driver's gaze to detect both visual and cognitive distraction. We trained our proposed method using two self-created datasets, then tested it in complex driving environments, such as stop sign-controlled intersections. The results show that using both eye movement and driver gaze direction improve detection accuracy and indicate the importance of dealing with each kind of distraction separately.**

## 1. Introduction

Despite the development and commercialization of advanced driver assistance systems, it is reported that 1.19 million people died in traffic accidents in 2021 (WHO, 2023). NHTSA (2018) found that 94% of critical, pre-crash events were the fault of human drivers, so decreasing human error would allow us to sharply decrease the number of traffic fatalities. Among the human factors that lead to serious traffic accidents is distracted driving, however driver distraction detection has proven to be challenging, since there are three kinds of distraction caused by very different factors; visual distraction, cognitive distraction, and manual distraction (NHTSA, 2010). But appropriate gaze behavior also depends on the driving environment. For example, in complex environments, such as stop sign-controlled intersections, drivers must detect many driving-related visual targets, requiring them to frequently scan wide areas (Figure 1).

Huang and Fu (2022) proposed a driver distraction detector using the Driver's Focus of Attention (DFoA). But it is impractical to detect driver distraction using a benchmark DFoA generated using the gaze behavior of multiple drivers. And the authors do not mention variation in eye movement, so their method may be inappropriate for complex environments.

In this study, we propose a distraction detector that can be applied to complex environments. It is considering both the DFoA and driver eye movement, and allowing us to detect two kinds of distraction, visual and cognitive. We do not deal with manual distraction in this study because we believe it would be more effective to use hand movement and facial orientation to detect it, rather than visual behavior. We train driver attention predictor using our own 'intersection driving dataset', so that we could use it in more complex environments such as stop sign-controlled intersections. Also,

*Fig. 1. Left: Dashcam image. Right: Driver's Focus of Attention (DFoA) as predicted when using the driver attention predictor trained with conventional dataset. While it can detect the pedestrian, it should also consider traffic signs, blind spots, etc. in this complex environment.*

we use our 'distracted gaze dataset' to detect both types of distractions.

## 2. Method

We used driver gaze direction and eye movement as inputs, and employed a simple perceptron as our distraction detector, following Huang and Fu (2022). Figure 2 shows the architecture of the proposed method.

### *2.1 Use of Driver's Attention Prediction*

Driver attention prediction is the task of predicting the DFoA. We consider the predicted DFoA to be the correct focus of the driver's attention for a particular driving situation, and we use the relationship between the predicted DFoA and the driver's eye movement as indicators of the distracted driving. However, driver attention prediction methods trained on conventional datasets are strongly influenced by other traffic participants such as pedestrians and automobiles, and are difficult to predict appropriate driver attention in complex environments such as stop sign-controlled intersections.

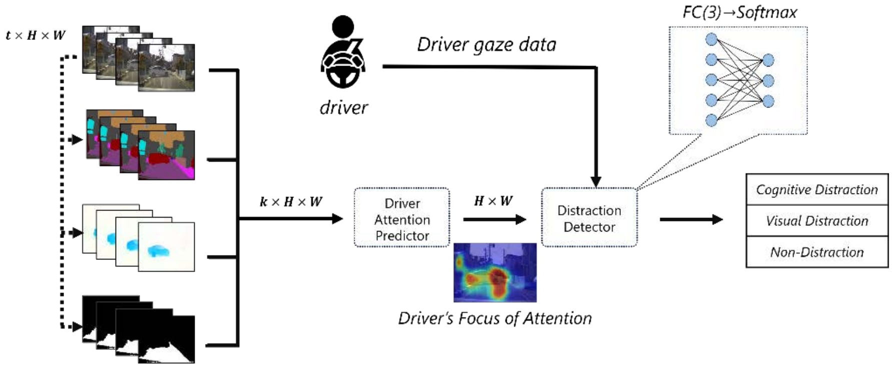



***Fig. 2.*** *Architecture of our Driver Distraction Detector.*

Therefore, we created the 'intersection driving dataset', which includes dashcam videos of driving in urban areas, to train a Gate-DAP (Zhao et al., 2023). After calculating the appropriate DFoA using the driver attention predictor, we acquired DFoA values based on the driver's actual gaze direction and the maximum values within a 5° diameter circle area centered on that gaze direction. This mimics the human central vision (Adam et al., 2009). The averages of these values for the last $T_{att}$ frames were used as distraction indicators that take the driving situation into account. At time $t$, indicator $pr_{avg,t}$ and $pr_{max,t}$ are expressed as shown in Equations (1) and (2):

$$pr_{avg,t} = \frac{1}{T_{att}} \sum_{t-T_{att}<k}^{t} a_{k,(i_k,j_k)} \tag{1}$$

$$pr_{max,t} = \frac{1}{T_{att}} \sum_{t-T_{att}<k}^{t} \max\left(a_{k,(i_k,j_k)} \in C_{(i_k,j_k)}\right), \tag{2}$$

where $(x_t, y_t)$ is the point on which the driver's gaze is focused, $a_{t,(x_t,y_t)}$ is the predicted DFoA value for point $(x_t, y_t)$, and $C_{(x_t,y_t)}$ is the inside of a circular area with a diameter of 5° centered at the point $(x_t, y_t)$.

### *2.2 Use of Eye Movement*

Percent Road Center (PRC) represents the percentage of driver's gaze on the center of the road, however it is not defined in detail (Khan & Lee, 2019). In this study, we use the center of the screen as the road's center, and set the road's central area to be a circle with a radius of 8 degrees, as in Kircher et al. (2009). We also employ standard deviations of eye movement in the width and height directions, giving us three eye movement indicators for point in time $t$; $mr_{prc,t}$ , $mr_{stdx,t}$ and $mr_{stdy,t}$ , respectively.

### *2.3 Distraction Detector*

We used a simple perceptron as our distraction detector, following Huang and Fu (2022). It is consisted of five inputs, $pr_{avg,t}$ , $pr_{max,t}$ , $mr_{prc,t}$ , $mr_{stdx,t}$ and $mr_{stdy,t}$ , and three outputs representing visual distraction, cognitive distraction, and non-distraction.

## 3. Experiment

### *3.1 Dataset*

We created two datasets for model training. One is the 'intersection driving dataset' consisting of video of driving in urban areas, which we used to train the attention predictor. To create the dataset, we first drove through stop sign-controlled intersections in Aichi, Japan and obtained video data for 130 intersections. After that, six subjects were asked to view these videos in a laboratory, as in Fang et al. (2022), and 12 attention fixation points were identified for each frame on average.

By having the participants in our experiment watch the 'intersection driving dataset' video under three different conditions, we also created 'distracted gaze dataset' for three eye gaze patterns associated with non-distraction, visual distraction, and cognitive distraction while driving. For the non-distraction condition, subjects viewed the video normally. For the visual distraction condition, we asked subjects to "look at the most interesting area of each scene, instead of focusing on the driving task". For the cognitive distraction condition, we gave the subjects a continuous n-back task (Kirchner, 1958). All six subjects watched the intersection driving video under all three conditions, then watched it again under the non-distraction condition, resulting in 3,080 scenes of gaze data. In this study, we used Tobii Pro Fusion set to 120fps as our eye tracker for both datasets.

### *3.2 Experimental Methods and Environment*

As the backbone of our DFoA prediction method, we employed a Gate-DAP (Zhao et al., 2023) trained with our 'intersection driving dataset'.

For distraction detection, the five indicators described in Sec. 2.3 were input to a simple perceptron, which classified the input into three states. All five inputs were calculated from all the video frames of each scenario. We used weighted cross-entropy as a loss function.

## 4. Results and Discussion

### *4.1 Driver's Attention Prediction*

Figure 3 shows a qualitative comparison of predicted DFoA trained with the DADA-2000 dataset and 'intersection

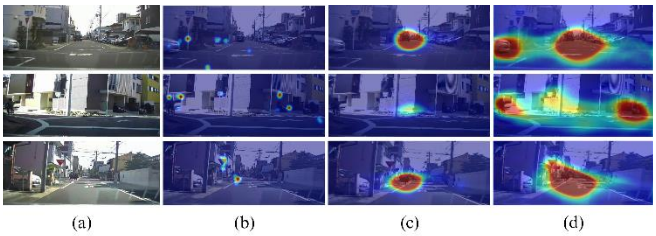



*Fig. 3. RGB frames(a), Ground-Truth(b), DFoA predicted driver attention predictor trained with the DADA-2000 dataset(c), our intersection driving dataset(d).*

driving dataset'. The DFoA trained with our dataset tends to pay more attention to areas containing road features and is approaching closer to the human-like driver's attention area.

### *4.2 Driver's Distraction Detection*

Table 1 shows classification accuracy, recall, and precision of the proposed method and two methods with only one of the two types of inputs, and Figure 4 shows a confusion matrix for the classification results of the proposed method. We can see that use of driver gaze direction and eye movement both contributed to improving distraction detection accuracy.

We further considered the two kinds of distractions as one class and evaluated the validity of classifying distractions by each state. Table 2 shows the results. The results showed that the accuracy decreases when two types of distraction states are included, indicating the importance of dealing with distraction separately for each state.

***Table 1.*** *Evaluation of effectiveness of driver distraction indicators. Gaze direction and DFoA consists of* $pr_{avg,t}$ *,* $pr_{max,t}$*, Eye movement consists of* $mr_{prc,t}$ *,* $mr_{stdx,t}$ *,and* $mr_{stdy,t}$

| Gaze direction and DFoA | Eye movement | Accuracy | Recall | Precision |
|---|---|---|---|---|
| ✓ | | 0.560 | 0.546 | 0.514 |
| | ✓ | 0.554 | 0.530 | 0.521 |
| ✓ | ✓ | **0.662** | **0.638** | **0.605** |

## 5. Conclusions

We proposed a distraction detector that uses both eye movement and direction of the driver's gaze to detect both visual and cognitive distraction and build it using two different datasets. Experimental results revealed that the driver's attention prediction model trained on the 'intersection driving dataset' tended to focus more on blind spots. Also, the direction of gaze contributed to improving the accuracy of both visual and cognitive distraction detection. Treating two types of distracted states as one type of distraction decreases classification accuracy, indicating that it is important to deal with each kind of distraction separately.

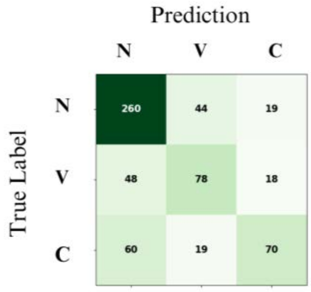



***Fig. 4.*** *Confusion matrix for the results of proposed method. N, V, and C represent Non-Distraction, Visual, and Cognitive, respectively.*

***Table 2.*** *Evaluation of two-class classification.*

| Target1 | Target2 | Accuracy | Recall | Precision |
|---|---|---|---|---|
| Non-distraction | Visual | 0.794 | 0.793 | 0.711 |
| Non-distraction | Cognitive | 0.795 | 0.793 | 0.713 |
| Non-distraction | Cognitive & Visual | 0.718 | 0.724 | 0.712 |

## References

Adam M. Larson and Lester C. Loschky. (2009). *The contributions of central versus peripheral vision to scene gist recognition*. *Journal of Vision*, Vol. 9, No. 10, pp. 6–6, 09

Fang, J., Yan, D., Qiao, J., Xue, J., and Yu, H. (2022). *DADA: Driver Attention Prediction in Driving Accident Scenarios. IEEE Transactions on Intelligent Transportation Systems*, 23(6), 4959-4971. doi: 10.1109/TITS.2020.3044678

Huang, T. and Fu, R. (2022). *Driver Distraction Detection Based on the True Driver's Focus of Attention*. *IEEE Transactions on Intelligent Transportation Systems*, 23(10), 19374-19386. doi: 10.1109/TITS.2022.3166208

Khan, M. Q. and Lee, S. (2019). *Gaze and eye tracking: Techniques and applications in ADAS*. *Sensors*, 19, 5540. doi: 10.3390/s19245540

Kircher, K., Ahlström, C. and Kircher, A. (2009). *Comparison of two eye-gaze based real-time driver distraction detection algorithms in a small-scale field operational test. Driving Assessment Conference* 5, 16-23.doi: 10.17077/drivingassessment.1297

Kirchner, W. K. (1958). *Age Differences in Short-Term Retention of Rapidly Changing Information*. *Journal of experimental psychology*, 55(4) 352-358. doi: 10.1037/h0043688.

NHTSA. (2010). *Overview of the National Highway Traffic Safety Administration's Driver Distraction Program*. https://www.nhtsa.gov/sites/nhtsa.gov/files/811299.pdf

NHTSA. (2018). *Traffic safety facts: Critical reasons for crashes investigated in the national motor vehicle crash causation survey.* https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812506

WHO. (2023). *Global status report on road safety 2023*. https://www.who.int/publications/i/item/9789240086517

Zhao, T., Bai, X., Fang, J. and Xue, J. (2023). *Gated Driver Attention Predictor. IEEE International Conference on Intelligent Transportation Systems.* doi: 10.1109/ITSC57777.2023.10422543

# Safe use of general controls: towards a Euro NCAP assessment protocol to target distraction by design

Ilse Harms [1*], Greta Große [2], Tina Gehlert [2], Florian Savona [3], Ruggero Ceci [4], Adriano Palao [5]

[1] Department of Applied Innovation, RDW, Europaweg 205, Zoetermeer, The Netherlands
[2] German Insurers Accident Research (Unfallforschung der Versicherer / UDV), Wilhelmstraße 43 /43 G, 10117 Berlin, Germany
[3] UTAC Montlhéry, Avenue Georges Boillot, 91310 Linas-Montlhéry, France
[4] Planning department, Trafikverket, Roda Vagen 1, 781 89, Borlange, Sweden
[5] European New Car Assessment Programme (Euro NCAP), Mgr. Ladeuzeplein 10, Leuven, Belgium
[*] iharms@rdw.nl

**Abstract: Being distracted while driving is a main contributor to traffic accidents. It is recognised that poor HMI design can lead to unnecessary long glances off the road view while performing a simple driving or non-driving related task. As part of the 2026 Rating Scheme, Euro NCAP will introduce the first-ever assessment procedure for General Driving Controls. Car manufacturers will thus be encouraged to implement good HMI practices to prevent visual-manual distraction by design. The framework will objectively assess both how a function is operated (e.g. by a direct physical input through a button, stalk, switch, versus by a touch display), as well as the quality of its implementation. A list of over sixty actions/tasks, e.g. operating the radio, windows or windshield wipers, will be distributed across six main categories of functions: 1) Hazard; 2) Driving; 3) Lane support systems; 4) Speed assistance; 5) Comfort; and 6) IVIS. Whether an HMI design is considered too distracting also depends on what exactly the driver tries to operate: is it about tuning the radio, changing the climate control, activating the fog lights, or triggering the hazard warning lights? Based on existing literature a checklist has been devised which targets poor HMI implementations.**

## 1. Introduction

It is well-known that driver distraction is a main contributor to traffic accidents. Modern vehicles contain an increasing number of communication, comfort and driving assistance systems, for which operation through touchscreens appear to have become a new standard (UDV, 2023). If the Human Machine Interfaces (HMI) are designed poorly, drivers need to take off their gaze from the road. This may cause unnecessary visual-manual distraction. A recent Swedish study pointed out that the HMIs of new cars are increasingly difficult to operate. They tested easy tasks such as changing the radio station or adjusting the climate control in various cars and found that task performance time was long and eyes were off the road for prolonged periods of time for cars using touchscreens instead of physical buttons (Vikström, 2022). The fact that these cars are available on the European market, is a strong indicator that current vehicle regulations – such as UN R121, including its 01 series of amendments (UNECE, 2023) – are not yet sufficiently covering this issue.

In Europe, vehicle type approval authorities, such as RDW, monitor the lower limit of which vehicles are admitted to the European market. Euro NCAP on the other hand, makes it clear to European consumers which vehicle models distinguish themselves in a positive or negative way in terms of safety. This is done on the basis of a rating system with stars: the more stars, the better. As a result, Euro NCAP also has a significant influence on the automotive sector.

Euro NCAP supports Safe Driving. If a consumer buys a 5-star Euro NCAP rated car, they must be able to trust it can be operated safely (Euro NCAP, 2022). Therefore, to avoid distraction by design, the aim is to draft a protocol to assess the safe use of general controls.

## 2. Method

### *2.1 Procedure*

The assessment protocol is being drafted in Euro NCAP's HMI & Human Factors Working Group, which consists of experts within the field of human factors and HMI, and/or experts on drafting and working with Euro NCAP assessment protocols. During the drafting process, feedback from the automotive industry was considered.

To develop the HMI assessment protocol for general driving controls, existing literature was reviewed. This was done to recognise existing best practice instead of producing a set of disruptive requirements or methodologies on a short notice, as the assessment will enter into force in January 2026.

### *2.2 Materials*

#### *2.2.1 The protocol format*

Euro NCAP assessments are performed by a certified inspector who follows the assessment protocol. Euro NCAP is familiar with different formats for assessment protocols, such as dossiers – in which the car manufacturer supplies the relevant documents and information to be assessed –, checklists with gradual ratings, and checklists based on pass/fail criteria. For the current assessment protocol the format of a checklist consisting of pass/fail criteria has been chosen. This, to reduce the burden on car manufacturers and meeting time and cost limitations of the test programme. Furthermore, the typical Euro NCAP assessment requirements had to be considered: repeatability and reproducibility. A test should produce the same outcome regardless of where it is assessed, who assesses it and how many times it is assessed.

#### 2.2.2 *Categories of functions*

The assessment protocol targets a broad range of controls. Drivers may use different functions while driving their car: from using the wipers, to changing the climate control, to tuning the radio. Therefore six main categories of functions have been defined. Each main category consists of various functions, as displayed in Table 1.

***Table 1****. Overview of the functions per main category*

| Main categories | Functions |
|---|---|
| 1) Hazard | Hazard light; e-Call; Horn |
| 2) Driving | Direction indicator; Vehicle settings; Windshield wipers (front and rear); Demisting (front, rear and side mirrors); Lights (interior, exterior, full/high beam, headlight height adjustment, fog lights [front and rear]); Screen brightness; Mirrors (side [external] and rear view); Sun visor |
| 3) Lane support systems | Lane centering; LKA |
| 4) Speed assistance | ACC; ISL |
| 5) Comfort | Climate controls; Windows; Sunroof; Seat adjustment |
| 6) IVIS | Audio entertainment; Calling and dialling; Text messaging; Navigation system; Vehicle status |

### 2.3 *Measures*

The current protocol needs to assess 1) function operation (e.g. by a direct physical input versus by a touch display), and 2) implementation quality.

The main building block used to assess *function operation* is based on a decision tree developed by UDV (2023), as displayed in Figure 1. This was used to define the minimally accepted implementation of the respective function (for detailed information about the methodology and background of the decision tree see UDV, 2023).

During assessment development for *implementation quality*, it was found that most literature on the safe use of in-vehicle controls concentrated on in-vehicle information systems (i.e., IVIS). Therefore the well-established NHTSA guidelines (2013) for IVIS have been revised, updated and adapted to be suitable for the assessment of multiple functions.

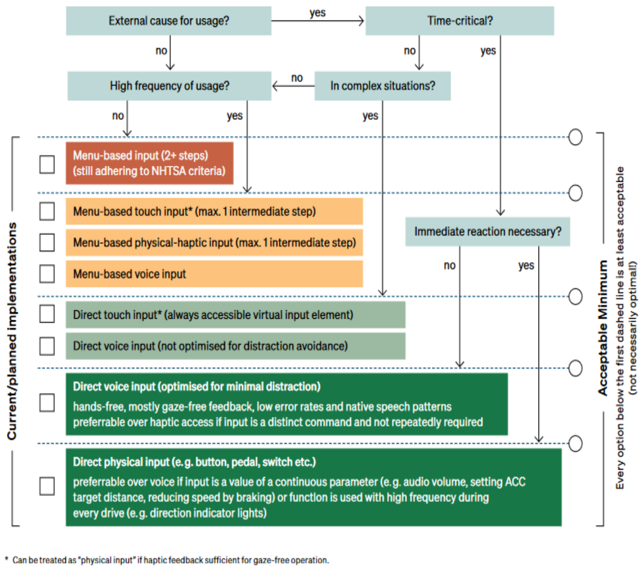



***Figure 1****. UDV's (2023) decision tree for function implementation: 'the minimally accepted implementation of the respective function can be read off from the conditions of use on the right-hand side, while the current / planned implementation of the function in the vehicle is indicated on the left-hand side'. Note that for the 2026 protocol this decision tree has been adapted to exclude voice commands.*

## 3. Results

### 3.1 *Draft checklist 2026*

The resulting draft checklist consists of a matrix design. On the vertical axis, it displays all various functions, grouped per category (Hazard, Driving, Lane support, Speed assistance, Comfort, IVIS). For each function one or various specific tasks have been defined. Some examples:

- to assess the function 'direction indicator' one task is to activate the left indicator,
- to assess the function 'climate controls' one of the tasks is to activate the air-conditioning,
- to assess the function 'audio entertainment' one of the tasks is to adjust the volume.

On the horizontal axis, per task, the minimally accepted implementation of the respective function is defined. This ranges from direct physical input i.e. producing kinaesthetic feedback for the user (e.g. buttons), to direct touch input i.e. always accessible but not producing kinaesthetic feedback, to menu-based physical haptic or touch input (max 1 intermediate step), till menu-based input (2+ steps). The accompanying pass/fail criteria resulted from following the UDV (2023) decision tree and are based on expert opinion and consensus within the Working Group.

Furthermore, per task, the quality of the function implementation is assessed, divided by function-related criteria and additional criteria for multi-step function tasks only. Criteria for both are displayed in Table 2.

***Table 2**. Overview of the function-related criteria and the criteria for multi-step function tasks only*

| Function-related criteria | Multi-step function tasks only criteria |
|---|---|
| allowing to have at least one hand on the steering wheel | not requiring an uninterrupted input sequence |
| legibility of presented texts | the possibility to resume an operator-interrupted sequence |
| timely and clearly perceptible system responses following driver input | avoiding automatic system-initiated loss of partial driver input |
| or at least providing a clearly perceptible indication in case system response time exceeds 2s | implementation of commands that erase driver inputs |
| production of any corresponding system status changes following a driver input | visually displaying previously entered data or current feature state |

### *3.2 Out of scope for the draft checklist 2026*

Although the GDV decision tree allows for taking voice input as an implementation into consideration, it is not covered by the draft assessment protocol for 2026. This is because currently there is no sufficient base in literature that would allow for defining the required criteria to assess the quality of voice-based input. For example it would require criteria for commands spoken in different accents and criteria for various wording which are ecologically valid commands to operate specific functions.

## 4. Discussion

Given the fact that current vehicle regulations are not providing a solid base to prevent distracting HMI designs, Euro NCAP is once again filling this gap. To that end, the current HMI assessment protocol is in its first implementation targeting designs prone to distraction and gives recognition to designs adhering to good HMI practices.

Its checklist will be evaluated during tests in a near future.

Furthermore, every 3 years, Euro NCAP updates requirements across various assessment protocols, continuously encouraging the automotive industry to make ever-safer cars. Therefore the current protocol is a dynamic document, further developed and validated by international studies, methods and guidelines. Next steps could be including the assessment of voice control and expanding assessment criteria towards intuitive design. The current 2026 HMI protocol is a first step to take HMI into account as part of Euro NCAP's rating scheme, using a dedicated protocol.

## 5. Conclusions

Whether an HMI design is considered too distracting depends on what exactly the driver tries to operate: is it about tuning the radio, changing the climate control, activating the fog lights, or triggering the hazard warning lights? To quantify this, criteria have been defined per function, assessing how a function is operated (e.g. by a direct physical input through a button, stalk, switch, versus by a touch display) as well as its quality of implementation. The foundation of the current draft assessment protocol is a checklist based on existing literature, which targets the worst HMI implementations for each function.

This new HMI protocol is part of Euro NCAP's star-rating and will be effective by January 2026.


## 6. Acknowledgments

The authors wish to express their gratitude to all the experts in Euro NCAP's HMI and Human Factors Working Group who have helped drafting this General Driving Controls assessment protocol and for putting in major effort to make this a success. Special thanks to the Euro NCAP Board for believing in us! Together with manufacturers we are now further paving the way towards a human-centred approach to safe driving.


## References


Euro NCAP (2022). *Euro NCAP Vision 2030. A Safer Future for Mobility*. Euro NCAP: Leuven, Belgium.

UDV (2023). *Distraction due to vehicle operation*. Compact Accident Research Report No. 125. Unfallforschung der Versicherer (UDV – German Insurers Accident Research): Berlin, Germany. https://www.udv.de/udv-en/distraction-due-to-vehicle-operation-137110.

National Highway Traffic Safety Administration (NHTSA) (2023). Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices. *Federal Register, 79*, Technical report 81, p. 24818-24890. https://www.federalregister.gov/documents/2013/04/26/2013-09883/visual-manual-nhtsa-driver-distraction-guidelines-for-in-vehicle-electronic-devices. Last visited on 23 May 2023.

UNECE (2023). *UN Regulation No. 121 - Rev.2 - Identification of controls, tell-tales and indicators (01 series).* Last update: UN R121.01 suppl. 6; date of entry into force 24/09/2023.

Vikström, F. D. (2022). Physical buttons outperform touchscreens in new cars, test finds. *ViBilägare*. https://www.vibilagare.se/english/physical-buttons-outperform-touchscreens-new-cars-test-finds. Last visited on 19 March 2024.

# Applying European NCAP Driver State Monitoring Protocols to Heavy Vehicle Fleets: Prevalence of Distraction Alerts in Real-World Commercial Transport Operations

Jonny Kuo[1*], Megan Mulhall[1], Brook Shiferaw[1], Shiyan Yang[1], Michael G. Lenné[1], Michael Fitzharris[2], and Darren Wood[3]

[1] Seeing Machines Ltd, 80 Mildura St, Fyshwick ACT, Australia
[2] Monash University Accident Research Centre, Monash University VIC, Australia
[3] Ron Finemore Transport Services Pty Ltd, Wodonga VIC, Australia
[*]jonny.kuo@seeingmachines.com

**Abstract: The 2023 European New Car Assessment Program (Euro NCAP) Occupant Status Monitoring (OSM) protocols define distraction behaviours that Driver State Monitoring (DSM) systems must detect to score maximum safety points. Variations of these protocols will likely be applied to heavy vehicles, but the prevalence, relevance, and user-acceptance of these behaviours in heavy vehicles has yet to be examined. The current analysis applied the Euro NCAP OSM protocols on distraction to a previously collected naturalistic driving study of a real-world heavy vehicle fleet, comprising 1487 hours of driving data across 82 truck drivers. Using the proposed Euro NCAP protocol definition, Long Glance Away events were classified at a rate of 1 event per 1.04 hours. While the overall event rate was comparable to what has previously been reported in drivers of passenger vehicles, truck drivers spent more time looking at driving-related regions of the vehicle, such as the passenger- and driver-side mirrors. Distinguishing between driving related and non-driving related regions, as well as further investigation into the context of these events, is likely to be needed to further optimise the user acceptance and efficacy of potential alerts issued by Euro NCAP-compliant DSM.**

## 1. Introduction

In 2023, testing protocols for Occupant Status Monitoring (OSM) were published by the European New Car Assessment Program (Euro NCAP). These protocols outlined distraction behaviours that must be detected in drivers of passenger vehicles to earn maximum safety points (Euro NCAP, 2022). These protocols, or variants thereof, will likely be adapted for use in heavy vehicles. However, the prevalence of distraction behaviour and driver alerting frequency has yet to be examined in this cohort. This data is needed to make evidence-based decisions on whether these protocols can be implemented in heavy vehicles as-is, or whether they require modification to adapt to the unique environment of the truck cabin.

Developed in recognition of the increasing technical viability for camera-based OSM to protect drivers from the well-established road safety problem of driver distraction, the Euro NCAP OSM protocols operationalise two distinct categories of distraction behaviour: Long Glance Away (LGA), defined as glances >=3s away from the forward roadway; and Visual Attention Time Sharing (VATS), defined as a cumulative 10s of glances away from the forward roadway within a 30s window, and where intermittent glances to the forward roadway (if present) do not individually exceed 2s in duration (Euro NCAP, 2022). While the rationale for detecting LGA and VATS behaviour are firmly established in the road safety literature (Euro NCAP, 2022; Klauer et al., 2006; Tivesten et al., 2019), the extent to which these behaviours are applicable, relevant, or user-acceptable in the context of heavy vehicles has yet to be examined.

Significant differences exist between the driving task in heavy vehicles versus passenger vehicles. The implications of these differences for operationalising OSM within the heavy vehicle context needs to be understood. Examples of these differences include the additional regulatory requirements imposed on heavy vehicle drivers around hours of service (National Heavy Vehicle Regulator, 2018), different road environments where the driving task is likely to occur, as well as differences in the cabin geometry of the respective vehicle types.

A further distinction between heavy vehicle and passenger vehicle drivers exists in the early adoption of aftermarket camera-based OSM in the heavy vehicle space. Referred to as Fatigue and Distraction Detection Technologies (FDDTs) by Australian heavy vehicle regulators, there is significant interest and growing regulatory pressure promoting the uptake of the devices across heavy vehicle fleets (Higginson et al., 2019). The use of these devices has been demonstrated to decrease the occurrence of fatigue (Fitzharris et al., 2017). The prevalence and general support of FDDT use in heavy vehicle fleets serves as an interesting counterpoint to the often-increased public scrutiny on safety that operators face.

To establish an evidence base for the applicability, relevance, and user-acceptability of the Euro NCAP OSM protocols in heavy vehicles, the current analysis applied the protocols to an existing real-world heavy vehicle naturalistic driving dataset.

## 2. Method

### 2.1 Dataset

A naturalistic driving study was conducted with an operational trucking fleet as part of the Advanced Safe Truck

Concept project (a Cooperative Research Centre Projects-funded partnership program) (Lenné, 2018). Ethics approval for this phase of the project was obtained through the Monash University Human Research Ethics Committee. Prior to recruitment, a series of presentations outlining the project and participation requirements were delivered to drivers and fleet managers. Drivers were then independently approached to obtain informed consent.

In total, 10 Volvo prime-mover vehicles were fitted with Seeing Machines' automotive grade driver monitoring system (DMS). The driving performance of 82 consenting drivers while they carried out their normal shifts for a period of up to 6 months was tracked. The DMS comprised a driver-facing infrared camera, mounted on the dashboard above the centre console. Operating at 46fps, the DMS recorded continuously while vehicle ignition was switched on. The DMS tracked driver features including head and eye position, glance location, and glance intersection with attention regions within the vehicle cabin (e.g. centre console, instrument cluster etc). Additionally, video data of the forward roadway, a wider-angle video view of the vehicle cabin, as well as vehicle kinematic data were recorded. A total of 218 shifts ranging from 5 – 12 hours in duration were included for analysis (total hours: 1487).

### *2.2 Analysis*

The dataset was analysed for Euro NCAP distraction behaviours, specifically Long Glance Away events, operationalised as continuous off-road glances of >3s duration. To calculate representative event rate statistics, DMS data from when drivers were undertaking scheduled breaks from driving or when the vehicle was moving <10km/h were excluded from the analysis.

## 3. Results

1428 LGA events were classified across 82 participants, with per-driver event ranges between 0.16 – 9.83 events/hour. On average, LGA events were classified once every 1.04 hours (0.96 events/hour).

To facilitate comparison with a reference passenger vehicle cohort, results from Mulhall et al.'s (2023) analysis of Euro NCAP distraction behaviours in car drivers (14 drivers across 167 hours) have been adapted with permission and presented in Table 1 below.

***Table 1.*** *LGA descriptive statistics and comparison between passenger vehicles and heavy vehicles. Passenger vehicle data adapted with permission from Mulhall et al. (2023)*

| | Passenger Vehicles | Heavy Vehicles |
|---|---|---|
| 1 event per x hours | 1.1 | 1.04 |
| Events per hour | 0.89 | 0.96 |
| Alert range | 0.07-4.55 | 0.16-9.83 |
| Non-driving related region proportion | 57.3% | 41.9% |
| Most frequent region | Console | Off road |
| Driver lap % | 8.6% | 7.3% |

Overall, LGA events were observed at a comparable rate across both populations. In the heavy vehicle dataset, LGA events were observed at a rate of 0.96 events/hour (1 event every 1.04 hours).

Descriptive statistics showed a difference in the distribution of LGA events across vehicle-interior attention regions between the two groups. However, as the cabin geometry (and, therefore, OSM attention regions) did not correspond exactly between the two groups, statistical comparisons were not conducted on specific attention regions. To illustrate these differences, a visualisation of the distribution of glances across the two cabin types is presented in Fig 1 below. (The relative location of glance regions in a passenger vehicle cabin can be found in Fig 1 of Mulhall et al. (2023)).

Compared with car drivers, truck drivers made a greater relative proportion of glances to the passenger side mirror, fewer glances to the centre console and instrument cluster, and more glances to 'off road other' (defined as any glance not to a defined off-road gaze region).

## 4. Discussion

The present analysis investigated the prevalence and characteristics of Euro NCAP-defined Long Glance Away (LGA) events in a real-world heavy vehicle fleet. Continuous, driver-facing OSM data from 82 truck drivers across 1487 hours of driving were analysed. Data from a previously published analysis of the prevalence of Euro NCAP LGA events in passenger vehicle drivers were used as a point of comparison.

During Euro NCAP-defined LGA events, heavy vehicle drivers distribute the location of their glances differently compared to passenger vehicle drivers, specifically, A greater proportion of glances by heavy vehicle drivers were made to the passenger side mirror, while a lower proportion of glances were made to other defined regions, such as the centre console. Despite these differences, overall event rates for LGA did not differ between heavy vehicle drivers and passenger vehicle drivers. Taken together, this suggests a need to distinguish between driving-related and non-driving-related regions when classifying LGA events as one way of optimising user acceptance of potential alerts.

However, this is not to conclude that LGA events to driving-related regions are necessarily safe and not in need of intervention (a gap in current research previously discussed in Mulhall et al. (2023)). Rather, we argue that there is a need to better understand the context in which these events occur, the drivers' intent in engaging with these behaviours, and then understanding and managing any potential safety impact that may arise. In the context of extending the current analysis, video data from the forward roadway, vehicle cabin (beyond the field of view of an OSM camera), as well as vehicle performance data could all be utilised to further the investigation.

## 5. Conclusions

Euro NCAP defined distraction events occur in real-world heavy vehicle driving, with LGA events being observed once every 1.04 hours. While the overall event rate


***Fig 1**. Number of long distraction events with >3s threshold for passenger vehicles (**A**) and heavy vehicles (**B**). Circle area is proportional to number of events. Off road other event location includes all glances not to a defined off road gaze region. (**A**) adapted with permission from Mulhall et al. (2023). To better facilitate comparison with (**A**), (**B**) has been mirrored from an image of a left-hand-drive cabin. As such, minor details (e.g. windscreen wiper orientation) will be incorrect. For illustrating attention regions, the image is otherwise representative.*

between passenger vehicle drivers and heavy vehicle drivers is comparable, the distribution of glance locations differs between the two cohorts, with truck drivers spending proportionally more time glancing at the passenger side mirror and less time at defined non-driving related regions such as the centre console. Clearly defining and distinguishing between driving related and non-driving related regions, as well as further investigation into the context of these events, is needed to optimise the user acceptance and efficacy of potential alerts, and in turn, to maximise the safety potential of OSM,

## 6. Acknowledgments

We acknowledge the funding received from the Commonwealth of Australia through the Cooperative Research Centre - Projects scheme and the cash and/or in-kind support provided by Seeing Machines, Monash University, Ron Finemore Transport and Volvo Group Australia. We also thank the extended team supporting this project: Fivaz Buys, Kyle Blay, Shafiul Azam, Gemma Hooper, David Moorhouse, Jeremy da Cruz, Tim Edwards and Rod Stewart at Seeing Machines; Timothy Horberry, Nebojsa Tomasevic, Brendan Lawrence, and Raphaela Schnittker at Monash University.

**References**

Euro NCAP. (2022). *Euro NCAP 2025 Roadmap: In Pursuit of Vision Zero*. Euro New Car Assessment Program.

Fitzharris, M., Liu, S., Stephens, A. N., & Lenné, M. G. (2017). The relative importance of real-time in-cab and external feedback in managing fatigue in real-world commercial transport operations. *Traffic Injury Prevention*, *18*(sup1), S71–S78. https://doi.org/10.1080/15389588.2017.1306855

Higginson, A., Dawson, D., & Sprajcer, M. (2019). *Research Report: Phase 2.Fatigue/Distraction detection technology use in the Australian road freight transport sector.* National Heavy Vehicle Regulator. https://www.nhvr.gov.au/files/202002-1120-phase-2-fatigue-distraction-detection-technology-australia.pdf

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data* (Technical Report DOT-HS-810594). National Highway Traffic Safety Administration. http://www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/Crash%20Avoidance/2006/DriverInattention.pdf

Lenne, M., Kuo, J., Fitzharris, M., Horberry, T., Mulvihill, C., Blay, K., Wood, D., & Peden, M. (2018). The Advanced Safe Truck Concept project: a partnership program developing future integrated driver monitoring technology. In A. Hržić (Ed.), *25th ITS World Congress, Copenhagen, Denmark, 17-21 September 2018* (pp. 1-8). ERTICO - ITS Europe.

Mulhall, M., Wilson, K., Yang, S., Kuo, J., Sletten, T., Anderson, C., Howard, M. E., Rajaratnam, S., Magee, M., Collins, A., & Lenné, M. G. (2023). European NCAP Driver State Monitoring Protocols: Prevalence of Distraction in Naturalistic Driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 00187208231194543. https://doi.org/10.1177/00187208231194543

National Heavy Vehicle Regulator. (2018). *Heavy Vehicle (Fatigue Management) National Regulation*. https://www.legislation.qld.gov.au/view/whole/html/inforce/current/sl-2013-0078

Tivesten, E., Victor, T. W., Gustavsson, P., Johansson, J., & Ljung Aust, M. (2019). Out-of-the-loop crash prediction: The automation expectation mismatch (AEM) algorithm. *IET Intelligent Transport Systems*, *13*(8), 1231–1240. https://doi.org/10.1049/iet-its.2018.5555

# Exploring Driver's self-report and observer ratings of driver drowsiness based on real road driving.

Aimé Tiadi[1*], Mohamed-Zaki Chellali[1], Florian Prince[2], Pascal Un[2], Pascal Quentin[1], Jean-Marc Tissot [1]

[1] Valeo Comfort and Driving Assistance, 6 rue Daniel Costantini, Créteil, France.

[2] Human Design Group, 7 ter Rue de la Porte de Buc, Versailles, France

[*]Corresponding Author: aime.tiadi@valeo.com

**Abstract: Several studies have been conducted since some decades to understand the driver drowsiness, its assessment and the ways to mitigate its effects. The driver self-report is a subjective measure well admitted in the drowsiness context. However, studies that compare it to the observer ratings in the Driver Monitoring System validation purposes are scarce.**

**This study compares driver self-reports using Karolinska Sleepiness Scale (KSS) and trained observer ratings in order to build a drowsiness system validation database. Fifty subjects were included in this study and drove with and without sleep deprivation on a real road environment. The subjects wore an Electroencephalograph (E.E.G) device for objective measures. Six independent raters were trained by human factor experts and assessed the driver drowsiness by analysing the videos.**

**The results showed significant differences between driving with and without sleep deprivation for both observer ratings and driver self-reports. Our study confirmed that drivers are capable to estimate their drowsiness state. Also, the results showed that the observer had the ability to measure the sleep deprivation effects.**

**This study showed that observers tend to rate the drowsiness state closer to objective E.E.G data than drivers.**

**We suggest to add the observer ratings and objective data to the driver self-reports as complementary measures to strengthen the drowsiness monitoring system validation database.**

## 1. Introduction

According to the European Road Safety Observatory (2018), 10 to 20% of crashes are due to drowsiness or fatigue. Consequently, the driver drowsiness is a key topic addressed in the driver road safety framework.

Driver's alertness and attention impairment is assessed traditionally by 3 means:

- **The neurophysiological assessment** (EEG, ECG, EMG, EDA) including brain waves, heart rate, skin conductance etc. (Anund et al., 2008; Sparrow et al., 2019; Hu & Lodewijks, 2020).
- **The Behavioural and performance assessment**: including eye tracking studies, ocular, head parameters, vehicle signals analysis (Wierwille et al, 1994; Friedrichs & Yang, 2010; Zhang et al., 2016); and psychomotor vigilance tasks (Lim & Dinges, 2008; Basner et al., 2011; Grant et al., 2017).
- **The Subjective assessment**: including the Karolinska Sleepiness Scale (KSS) rating with which the driver estimates his own alertness and sleepiness states (Akerstedt & Gillberg, 1990; Akerstedt et al., 2016). The KSS is widely used and the largest subjective measurement tool of drowsiness.

Driver's drowsiness is also assessed by observer ratings. Hanowski et al. (2000) carried out a study about the impact of local short haul operations on driver fatigue using naturalistic field data. Based on Observer Rating of Drowsiness (ORD) method (Wierwille et al., 1994), the authors reported video analyst drowsiness assessment using five-point scale without comparing to the driver's self-reports. The results showed that video analysis is able to assess different levels of fatigue. A driving simulator study about the driver's awareness of sleepiness showed that the drivers are aware enough to their sleepiness level leading them to avoid hazards and crash risk (Williamson et al., 2014). The authors measured this awareness using drivers' self-reports only.

Our real road study, aims to compare driver self-reports using Karolinska Sleepiness Scale (KSS) and trained observer ratings in order to address the following questions:

- Are driver self-reports sufficient to assess sleepiness and build accordingly a database leading to validate a system that monitor driver drowsiness?
- Do observer ratings provide additional values to strengthen driver drowsiness assessment and robustify consequently the system validation database?

## 2. Method

### *2.1 Participants*

The study included 50 participants having valid driving licence with 50% male and 50% female. Their age ranged between 20–65 years old and more (average: 40.18 years; SD: 15.39), and they drove regularly.

The participants were recruited with the help of medical experts of sleep located in the south-west of France;

neither sleep disorders nor medical advice preventing participants from driving were mentioned and documented; if so they were excluded from the study.

All participants filled and signed a written informed consent document prior to experiments. This study was approved by the local Ethics Committee.

### *2.2 Procedure*

Each participant performs 2 driving sessions: one driving for baseline and another for drowsiness session. The baseline session, **condition A,** in which the participants are not deprived of sleep**.** The drowsiness session, **condition B,** in which participants have deprived of sleep.

The participants provided their self-estimations of sleepiness each 5 minutes during both conditions, using the 9 levels of Karolinska Sleepiness Scale (KSS), Åkerstedt T, Gillberg M (1990): KSS level 1 means "extremely alert" until KSS level 9 meaning "Very sleepy, great effort to keep awake, fighting sleep". E.E.G device was also used as ground truth. The **figure 1** summarizes the protocol.

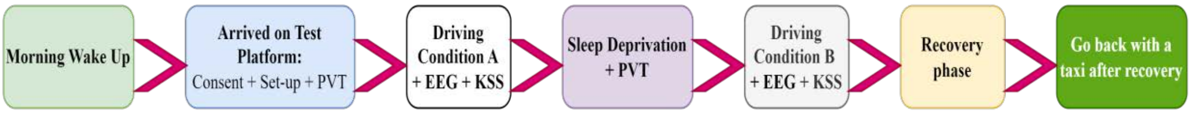



***Fig. 1.*** *Summary of the driving protocol*

### *2.3 Observer Ratings*

Six trained observers rated driver state by using observable drowsiness parameters (e.g: blink frequency, eye closure; yawning; movements on seat etc.) defined by human factor experts on the driver drowsiness topic. These parameters were matched to the KSS levels. For each driving session there were 3 independent observers who rated twice the same video. Video analysis are performed after the data collection. The average concordance rate of the observer judgements about drowsiness state is 0.92.

## 3. Results

### *3.1 Ratings In each Driving Condition*

The results showed that the mean observer ratings in driving condition B (with sleep deprivation) is higher than mean observer ratings in driving condition A (without sleep deprivation). Kruskal-wallis test showed significant differences between Condition A and Condition B for Observer ratings (H=682, 85; $P < 0.0000$). **Figure 2** depicts the mean observers' ratings in the both driving conditions.

The results showed that the mean driver self-reports in driving condition B is higher than mean driver KSS self-reports in driving condition A. Kruskal-wallis test showed significant differences between both conditions for drivers' self-reports (H=1047, 73; $p < 0.0000$). **Figure 3** depicts the mean KSS values of self-reports in the both driving conditions.

### *3.2 Differences Between Observer ratings and Driver Self-Reports*

The comparison of driver drowsiness estimation between observer ratings and self-reports showed that observer ratings are significantly different from driver self-reports in both driving conditions. Kruskal-wallis test showed: (H=252, 44; $p < 0.0000$) for condition A and (H=0, 11; $p < 0.0000$) for condition B.

In order to know if the observers over/under estimated or if the drivers over/under estimated the drowsiness level, we compare these two kind of ratings to objective E.E.G measures. In the condition A, there was no difference between observer ratings and driver self-reports as shown by the confusion matrix **(See figure 4**). However, in the condition B, the convergence percentages between E.E.G data and observer ratings are higher than E.E.G data and driver self-reports as shown by the confusion matrix **(See figure 5**). Least drowsy state from E.E.G outputs is estimated at 74, 51 % and 65, 38% as least drowsy state by observers and drivers respectively. Drowsy state from E.E.G outputs is estimated at 94, 12% and 87, 5% as drowsy state by observers and drivers respectively.

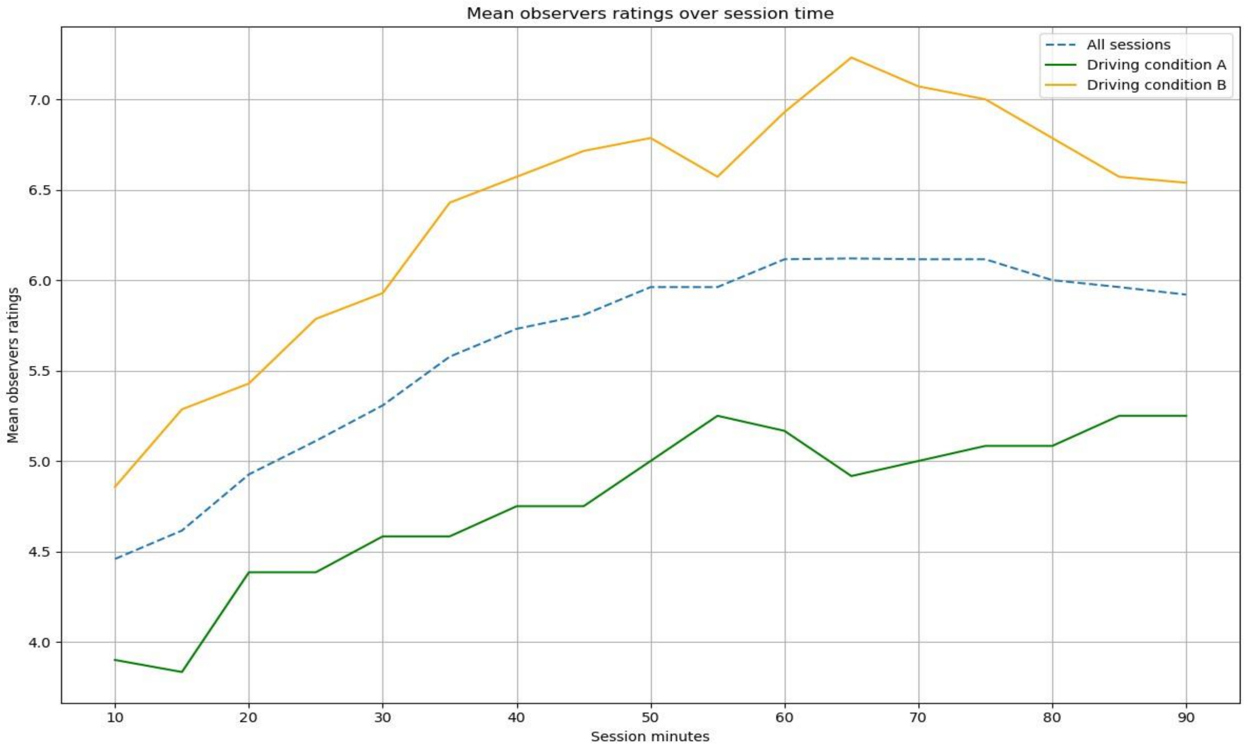



*Fig. 2.* *Mean KSS values of* ***Observers*** *ratings for driving condition without sleep deprivation (Condition A) and with sleep deprivation (Condition B).*

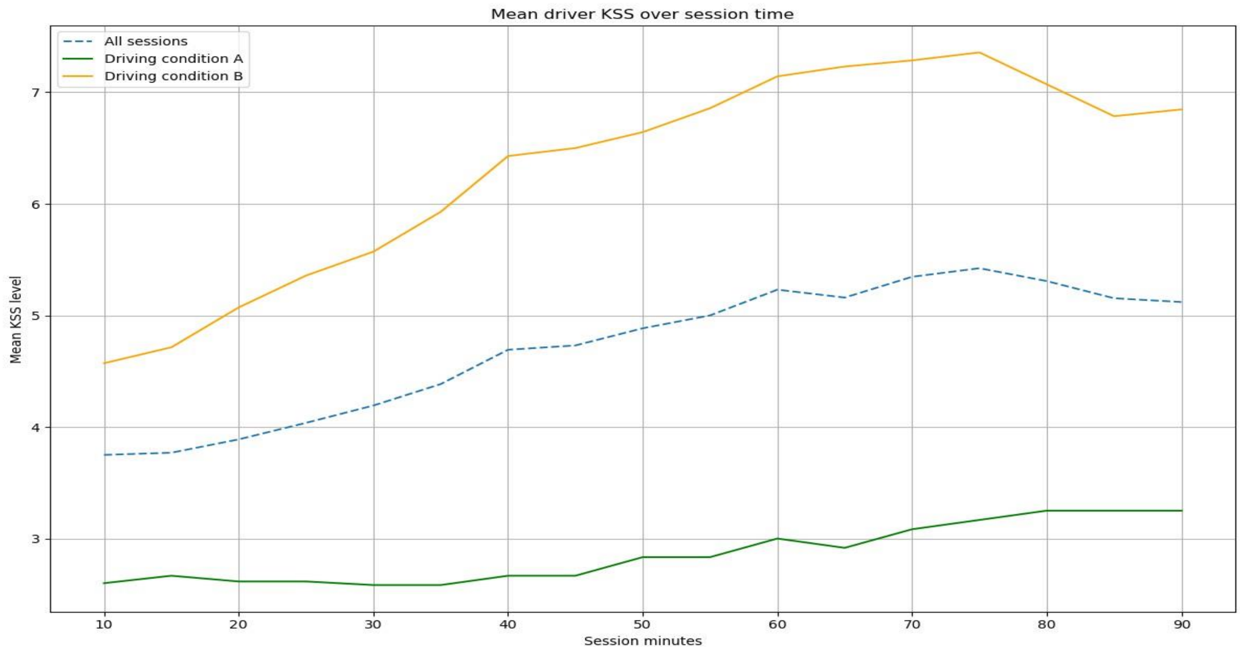



*Fig. 3.* *Mean KSS values of* ***driver's self-reports*** *for driving conditions without sleep deprivation (Condition A) and with sleep deprivation (Condition B).*

| Condition A | | Observer Ratings | |
|---|---|---|---|
| | | Not drowsy (KSS <7) | Least drowsy KSS (=7) |
| EEG Outputs | Not drowsy | 100.00% | 0.00% |
| | Least drowsy | 100.00% | 0.00% |

| Condition A | | Driver Self-Reports | |
|---|---|---|---|
| | | Not drowsy (KSS <7) | Least drowsy KSS (=7) |
| EEG Outputs | Not drowsy | 100.00% | 0.00% |
| | Least drowsy | 100.00% | 0.00% |

***Fig. 4.*** *Confusion Matrix showing drivers' state from E.E.G outputs (Objective data), Observer ratings and drivers' self-reports for driving condition without sleep deprivation (Condition A.)*

| Condition B | | Observer Ratings | |
|---|---|---|---|
| | | Least drowsy KSS (=7) | Drowsy (KSS >=8) |
| EEG Outputs | Least drowsy | 74.51% | 25.49% |
| | Drowsy | 5.88% | 94.12% |

| Condition B | | Driver Self-Reports | |
|---|---|---|---|
| | | Least drowsy KSS (=7) | Drowsy (KSS >=8) |
| EEG Outputs | Least drowsy | 65.38% | 34.62% |
| | Drowsy | 12.50% | 87.50% |

***Fig. 5.*** *Confusion Matrix showing drivers' state from E.E.G outputs (Objective data), Observer ratings and drivers' self-reports for driving condition with sleep deprivation (Condition B.)*

## 4. Discussion

Our Study confirmed that drivers are able to judge their drowsiness states as reported in previous researches (Williamson et al., 2014; Fors et al., 2016). Also, the observer ratings' approach was sensitive to the sleep deprivation effect and allowed to assess drowsiness state. Our findings about the differences between observer ratings and driver self-reports are in line with previous researches of Anund et al., (2013) and Ahlstrom et al. (2015). These studies reported low correspondence between observer ratings and driver self-reports. Using the E.E.G objective data we found that sometimes, both drivers and observers overestimated the least drowsy state while underestimated the drowsy state. But these over or under estimations are more important in driver self-reports compared with the observer ratings: E.E.G outputs least drowsy state is 34, 62% vs 25, 49% of drowsy state for drivers and observers respectively; E.E.G outputs drowsy state is 12, 5% vs 5, 88% of least drowsy state for drivers and observers respectively.

Maybe the driving environment could explain these differences; while Williamson et al. (2014) study was on driving simulator our study was on real road. Indeed, Fors et al. (2016) found that drivers rated higher KSS levels in the driving simulator compared to the real road. Also, in the naturalistic field study of Hanowski et al. (2000; 2011), the authors reported that driver did not perceive drowsiness as critical as it is in long-haul trucking.

Another explanation could be that observers and drivers do not measure the same elements. In our study, the observers' judgment are based on behavioural face and head parameters while drivers report their own feeling about drowsiness. This is why we suggest to use the observer ratings and E.E.G data as additional measures that could complete the drivers' subjective assessment in the context of driver monitoring system validation database.

## 5. Conclusion

The current study showed that both observer ratings and drivers self-reports are able to recognize the drowsiness state. Since these two methods have not the same assessment basis, we recommend to use them as complementary measures to build the drowsiness monitoring system validation database including objective data.

Further study is useful to confirm the relevance of this integrated validation database.

## 6. Acknowledgments

We would like to thank our university and medical partners. The authors are very thankful to University and University Hospital Center.

## References


Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neurosciences*, *52*(1-2), 29–37.

Åkerstedt, A.M., Kecklund, G., & Åkerstedt, T. (2016). Comparing two versions of the Karolinska Sleepiness Scale (KSS). *Sleep and Biological Rhythms*, *14*(3), 257-260.

Anna, A., Fors, C., David, H., Torbjörn, A. & Kecklund, G., 2013). Observer Rated Sleepiness and Real Road Driving: An Explorative Study. *PLoS ONE, 8*(5), 1-8.

Anund, A., Kecklund, G., Peters, B., Forsman, Å., Lowden, A. & Åkerstedt, T. (2008). Driver impairment at night and its relation to physiological sleepiness. *Scandinavian Journal of Work, Environment & Health, 34*(2), 142-150.

Arefnezhad, S., Samiee, S., Arno Eichberger, A., & Nahvi, A. (2019). Driver Drowsiness Detection Based on Steering Wheel Data Applying Adaptive Neuro-Fuzzy Feature Selection. *Sensors*, *19*(943), 1-14.

Basner, M. & Dinges, D.F. (2011). Maximizing Sensitivity of the Psychomotor Vigilance Test (PVT) to Sleep Loss. *Sleep*, *34*(5), 581-591.

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D. & Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience and Biobehavioral Reviews, (*44), 58-75.

Ahlstrom, C., Fors, C., Anna, A., & Hallvig, D. (2015). Video-based observer rated sleepiness versus self-reported subjective sleepiness in real road driving. *Eur. Transp. Res. Rev.*, *2015*(38), 1-9.

Friedrichs, F. & Yang, B. (2010). Drowsiness Monitoring by Steering and Lane Data Based Features Under Real Driving Conditions. *European Signal Processing Conference*, (18), 209-213.

Fors, C., Ahlstrom, C. & Anund, A. (2016). A comparison of driver sleepiness in the simulator and on the real road. *Journal of Transportation Safety and Security, 10*(1-2), 72-87.

Grant, D.A., Honn, K.A, Layton, M.E., Riedy, S.M. & Van Dongen, H.P-A. (2008). 3-minute smartphone-based and tablet-based psychomotor vigilance tests for the assessment of reduced alertness due to sleep deprivation. *Behavior research Methods, 49*(3), 1020-1029.

Gwak, J., Hirao, A. & Motoki Shino, M. (2020). An Investigation of Early Detection of Driver Drowsiness Using Ensemble Machine Learning Based on Hybrid Sensing. *Applied Sciences*, *10*(2890), 1-13.

Hallvig, D., Anund, A., Fors, C., Karlsson, J., G., Wahde, M. & Åkerstedt, T. (2013). Sleepy driving on the real road and in the simulator – A comparison. *Accident Analysis and Prevention*, (50), 44-50.

Hanowski, R. J., Wierwille, W. W., Garness, S. A., & Dingus, T. A. (2000). Impact of Local/Short Haul Operations on Driver Fatigue. *Federal Motor Carrier Safety Administration, U.S. Department of Transportation*. Report No. DOT-MC-00-203.

Hitchcock, E. M., & Neubauer, C. (2013). Active and passive fatigue in simulated drivin g: Discriminating styles of workload regulation and their safety impacts. *Journal of Experimental Psychology: Applied, 19*(4), 287–300.

Hu, X. & Lodewijks, G. (2020). Detecting fatigue in car drivers and aircraft pilots by using non-invasive measures: The value of differentiatio n of sleepiness and mental fatigue. *Journal of Safety Research, (*72), 173–187.

Lim, J. & Dinges, D.F. (2008). Sleep Deprivation and Vigilant Attention. *Annals of the New York Academy of Sciences*, *1129*(1), 305-322.

Muzet, A., Pébayle, T., Langrognet, J. and Otmani, S. (2003). AWAKE pilot studyno.2 : Testing steering grip sensor measures. Technical Report IST-2000- 28062, CEPA.

Saxby, D. J., Matthews, G., Warm, J. S., Sparrow, A.R., LaJambe, C.M.

& Van Dongen, H.P-A. (2019). Drowsiness measures for commercial motor vehicle operations. *Accident Analysis and Prevention, (*126), 146-149.

Wierwille, W.W., Ellsworth, L.A., Wreggit, S.S., Fairbanks, R.J., & Kirn, C.L. (1994). Research on Vehicle Based Driver Status/Performance Monitoring; Development, Validation, and Refinement of Algorithms For Detection of Driver Drowsiness. *National Highway Traffic Safety Administration*, *808* (247), 1-219.

Williamson, A., Rena, F., Jake, O. & Raphael, G. (2014). Are drivers aware of sleepiness and increasing crash risk while driving? *Accident Analysis and Prevention*, (70), 225-234.

# Technology-Driven Framework to Mitigate Driver Distraction

Pnina Gershon*

[1] Center for Transportation & Logistics, AgeLab, Massachusetts Institute of Technology, 1 Amherst Street, Cambridge, MA, USA
* pgershon@mit.edu

**Abstract:** The proliferation of in-vehicle systems, personal electronics, and advanced driver-assistive systems may exacerbate distraction, but at the same time it may also present an opportunity to utilize technology to mitigate distracted driving. This paper underscores the need to understand human-technology interactions and introduces a technology-driven framework consisting of three components: (i) Monitor, which provides the functionality related to sensing the driver state within the dynamic driving environment (guided by expected behaviors); (ii) Manage, which oversees decision-making and behavior changes by implementing interventions like multimodal alerts and restricting functionalities; and (iii) Motivate, which targets sustainable behavioral shifts by employing incentives, coercion, and gamification to foster positive changes. Through the synergy of the Monitor, Manage, and Motivate components, the 3M framework can guide further development of technologies, incorporating advanced sensing, personalized interventions, and tighter automation integration. This technology-driven framework strives to promote safe driving behaviors while maintaining desired levels of driver attention and introducing long-lasting changes in behavior.

## 1. Introduction

Distracted driving is a leading contributing factor to vehicle crashes in many countries [1]. According to the National Highway Traffic Safety Administration (NHTSA), in 2019 distracted driving contributed to 3,142 fatalities and an estimated additional 424,000 injuries on U.S. roadways [2]. While the burden associated with distracted driving due to loss of contextual awareness is likely to increase with the prospect of increasing vehicle automation, this can also be an opportunity to leverage new technological capabilities that reduce driver distraction and other risk behaviors.

## 2. Driving Automation and Distracted Driving

Consumer vehicles currently on the market are increasingly equipped with partial automation systems that can simultaneously control the longitudinal and lateral vehicle kinematics on a sustained basis. The driver remains responsible for monitoring the automation/environment and performing object/event detection, response selection, and execution. When using driving automation, the driving demands are lowered and the driver role pivots toward monitoring. Under these circumstances, it is challenging to maintain attention to the driving task and drivers often use the "freed-up" resources to do other things than driving. This tendency is amplified even more by the increased availability of portable electronics and in-vehicle technologies. These changes in driver behavior are observed although partial automation systems like Tesla Autopilot are delivered with formal statements that drivers should stay vigilant and monitor the road at all times, emphasizing that drivers are still subjected to the same legislation and laws related to driver distraction. The overall concerns regarding driver readiness or failures to take back control in time-critical situations due to distraction are underpinned by high-profile crashes of vehicles equipped with partial automation and NHTSA's standing general order requiring manufacturers to report crashes involving the use of automation [3].

## 3. Mitigating Driver Distraction

Many of the building blocks for the tools needed to mitigate driver distraction already exist or are in the making. However, mitigating distraction successfully requires more than creating the technology, building the infrastructure, promoting legislation, or launching media campaigns. It is necessary to bring all these pieces together into a framework that can explain and guide how to change driver behavior in a desirable way, that is grounded in theory and applicable to engineers, designers, and policymakers. In contrast to models that capture typical driver behavior or map the boundaries of performance, the proposed *Monitor*, *Manage*, and *Motivate* framework (the 3M framework hereafter) is a process model that methodically operationalizes changes in driver behavior. This framework can be used to generate insights and provide behavioral design recommendations to achieve sustainable mitigation of driver distraction.

## 4. The 3M Framework

The Monitor, Manage, and Motivate components in the 3M framework (Fig.1) are responsible for handling multiple aspects of the distracted driving problem. When combined and connected with the relevant information streams, these components result in a framework to promote and support safe driving by managing attention on a moment-to-moment basis and inducing long-lasting changes to driver motivations.

### *4.1 Monitor*

This component provides all the functionality related to sensing the driver state within the dynamic driving environment (guided by expected behaviors) and communicates the driver and environment state to the *Management* component. Driver monitoring systems (DMS) offer new opportunities to monitor the driver's state in real-time, including distraction. Driver monitoring can utilize a range of sensors, applicable for capturing many aspects of behavior and for tracking behaviors at different granularity levels. Based on technological advancements, ongoing research, and regulatory and rating agency requirements, DMSs are expected to become prevalent and a standard feature in new vehicles.

Additional sensor data can come from ongoing efforts to develop sensing capabilities that monitor the physiological state of the driver (e.g., impairment due to fatigue, alcohol, or THC) and characterize interactions with in-vehicle systems (e.g., frequency of tapping on the infotainment screen). Other types of sensors focus on capturing driving performance and kinematics. Measures like speeding, standard deviation of speed, and lane position have been used separately or in combination with data from other sensors to infer on driver's state within the driving environment.

Recent technological developments also allow for external monitoring of the driver state. For example, cameras mounted on traffic lights can capture driver activities within the vehicle and driving performance (e.g., speed and red-light running), thus delivering external monitoring capabilities [4]. In the 3M model, the monitoring component is guided by inputs of the expected or desired behavior of the driver and long-term expected behavioral patterns (from the *Manage* and *Motivate* components (Fig.1).

### *4.2 Manage*

This decision-making component is responsible for the execution of tactical changes in driver behavior. The successful operation of a management component depends on: (i) robust and mature monitoring that delivers reliable data; (ii) formalized rules and structured logic defining acceptable and unacceptable states and behaviors and, (iii) the availability of diverse and multimodal interventions. These capabilities are foundational for ongoing decision-making needed when managing immediate and short-term driver behavior.

Informed by the monitoring component this component takes a control theory approach for managing the gap between observed and desired behaviors. It maps the driver state onto constructs like visual attention, risk-taking propensity, automation use, misuse, and abuse. If the management component detected that the observed behavior deviates from the desired behavior threshold, it would issue an intervention to incite a behavior change. Thus, to direct driver behavior effectively, this component may rely on an ability to deliver interventions that are noticeable and can communicate the need for a timely response. Restriction-based prevention is another intervention to manage driver behavior that uses in-vehicle technology and/or smartphone applications to limit the opportunity to engage in distracting activities. For example, cellphone manufacturers offer the Do Not Disturb While Driving mode which targets distracted driving by prohibiting calls and texts and blocking audio features and specific applications when driving. Apple's CarPlay and Google's Android Auto take a different approach, rather than restricting functionality, these tools project the interaction from the smartphone to the car's infotainment displays, which are often larger than the phone screen.

### *4.3 Motivate*

To achieve significant and sustainable change in driver behavior there is a need for a coordinated set of interventions and supporting activities that target behavioral patterns. [5] investigated drivers' willingness to engage in distracted driving and found that decisions are strongly related to motivations and "lifestyle" perception. For example, reducing manual cellphone use while driving goes beyond the moment-to-moment management of driver attention and involves changes in intentions and how choices are made. Hence, the motivation component in the 3M model refers to the use of interventions and supporting activities to shape behaviors, habitual processes, emotional responses, and analytical decision-making [6]. Social norms, education, and awareness campaigns are some of the inputs that inform how to motivate drivers to engage in expected behavioral patterns. There are multiple types of interventions to shape driver motivations; some provide incentives that promote and reinforce desired behavior, while others demote or eliminate undesired behaviors through intimidation and punishments. Other relevant approaches include gamification, enablement, and scaffolding [7]. An example of an enablement-based intervention that encompasses behavioral support to promote safe driving is the simplification of Bluetooth pairing of cellphones with in-vehicle infotainment systems.

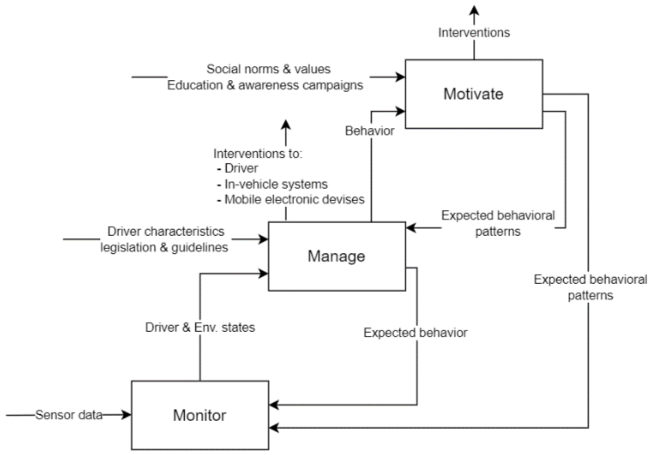



***Fig. 1.*** *The 3M framework core functional components and information flow.*

## 5. Conclusions

The increased functionality of in-vehicle systems and the prevalence of portable electronics, along with rapid developments of advanced driver-assistive systems (ADAS), dramatically increase the burden of distracted driving due to loss of contextual awareness. This study highlights the importance of understanding how humans interact with technology in both positive and negative ways and proposes a technology-driven framework to monitor, manage, and motivate driver behavior, prevent distraction, and mitigate its harmful effects.

Through the synergistic functioning of the Monitor, Manage, and Motivate components, the 3M framework holds the potential for further development and advances as technology continues to evolve. Future implementations of this framework might include driver attention management technologies with enhanced sensing and detection, personalized interventions, tighter integration of automation, and context-aware adaptation. Leveraging technologies, the framework endeavors to promote safe driving behaviors while maintaining desired levels of driver attention and introducing long-lasting changes in driver behavior.

## 6. Acknowledgments

I thank my colleagues, Dr. Lisa D'Ambrosio and Mr. Bruce Mehler, for their insightful feedback and comments that contributed to this work. The views and conclusions expressed are those of the author and have not been sponsored, approved, or endorsed by their affiliated organization.

**References**

[1] The World Health Organization. "Global status report on road safety," 2018. [Online]. Available: *https://www.who.int/violence_injury_ prevention/road_safety_status/2018/ en/*, Accessed: Jan. 15, 2020.

[2] National Center for Statistics and Analysis. (2022, May). Distracted Driving 2020 (Research Note. Report No. DOT HS 813 309). *National Highway Traffic Safety Administration*.

[3] National Highway Traffic Safety Administration. (2021). Standing General Order 2021-01: Incident Reporting for Automated Driving Systems and Level 2 Advanced Driver Assistance Systems [WWW Document]. *URL https://www. nhtsa. gov/sites/nhtsa. gov/files/2021-06/Standing_General_Order_2021_01-digital-06292021. pdf (accessed 8.1. 21).*

[4] Thompson, J. P., & Wundersitz, L. N. (2024). Mobile phone detection cameras in South Australia. *Prevention*, 119, 225-236.

[5] Lerner, N., Singer, J. P., & Huey, R. (2008). Driver strategies for engaging in distracting tasks using in-vehicle technologies (No. HS-DOT-810-919). United States. *National Highway Traffic Safety Administration*.

[6] Michie, S., Van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation science*, *6*(1), 1-12.

[7] Ebert, J., Xiong, A., Halpern, S., Winston, F. K., McDonald, C., Rosin, R., ... & Delgado, M. (2022). Summary Report: Comparative Effectiveness of Alternative Smartphone-Based Nudges to Reduce Cellphone Use While Driving (No. FHWA-HRT-22-057). *United States. Federal Highway Administration*. Office of Research, Development, and Technology.

# CogniDrive: Cognitive Perceptual Attention Factors Reduce Distraction

Paul Hemeren [1*], Julia Mattsson [2], Edin Pilipovic [3], Eric Moragues [4]

[1] Department of Informatics, University of Skövde, Skövde, Sweden
[2] Distancify AB, Science Park, Skövde, Sweden
[3] Tivala AB, Skövde, Sweden
[4] Autoliv AB, Vårgårda, Sweden
[*] paul.hemeren@his.se

**Abstract: As vehicles become increasingly automated, it is important to have a functioning inattention and attention collaboration between the driver and the semi-autonomous vehicle. In the case of semi-autonomous vehicles, the driver is not completely disengaged but still bears responsibility for driving. Since only certain functions are automated, the vehicle needs to be able to give the driver clear attentional feedback about an inattentional traffic situation and prompt the driver when he or she needs to avoid a traffic collision. Eight traffic warning modalities were used to investigate the movement from inattention to attention: haptic, visual, auditory, visual-auditory, haptic-visual, haptic-auditory, haptic-visual-auditory, and no-feedback. The results showed that the driver's reaction time was significantly shortest when the system provided haptic attention feedback as its own modality that interacted with the car drivers in a simulator. The visual factor was also used via previous research areas that provided drivers with visual information on the instrument panel. The visual and auditory information was provided via previous attention criteria. The other conditions that were slower than the haptic factor were the three different combinations with the haptic factor.**

## 1. Introduction

Driver attention direction is a clear relation to driver inattention in the sense that the levels of traffic severity vary where drivers need to move their attention in order to maintain a high traffic safety level.

Today, semi-autonomous vehicles are common on the roads, and levels of automation have to do with functions that were previously performed by humans as complete or partially replaced by a computer system. In the case of semi-autonomous vehicles, the driver is not completely disengaged but still bears responsibility for driving (European Commission, 2021; Favaro, Eurich & Rizvi, 2019; Gaffary, & Lécuye, 2019; Michelaraki et al., 2023; Masello, Sheehan, & Castignani, 2023). The arrival of such automation represents a steady development from the fully manual vehicles we have today, to the fully automatic vehicles we have in the future (Casner, Hutchins & Norman, 2016).

Studies have shown that most drivers react in 1-2 seconds for high-priority warnings such as frontal collision warnings (ITS Informal Group, 2011). Many common modalities in today's vehicles use visual and auditory feedback for ADAS-warnings such as TOR (take-over request), ACW (adaptive cruise warning), and FCW (forward collision warning). Forenbrock et. al (hereafter mentioned as NHTSA) made a study in 2011 where they tested different combinations of feedback in frontal collision warnings -including haptic seat belt feedback. Results showed that haptic feedback was involved when drivers reacted the fastest. The haptic feedback as a standalone modality, however, showed the longest reaction time. The shortest time was with the combinations of haptic, visual and auditory feedback.

This research classifies driver reaction times according to a standardized cognitive attention metric that can be utilized for assessing safety contexts, which are then used to dictate adaptive human machine interface (HMI) functions in real-time (Riener, Jeon, Alvarez & Frison, 2017; Xin et al., 2021). The major contribution is this relation between cognitive attention factors, driver monitoring, and safety levels. An anti-collision warning system (advanced driver assistance system, ADAS) was used to provide warnings to prompt the driver to act and avoid an accident. Three different high priority warning system modalities in frontal collision were used: haptic, visual, and auditory. These modalities were based on retrieved literature and a participant survey in a driving simulator. The effect of the seat belt haptic feedback on the driver's reaction time in frontal collision warning was tested as its own modality as well as in combinations with visual and auditory feedback.

In addition, this research presents the concept of driving analytics, which further develops the AI-based data analysis methodologies to enhance the accuracy of context detection for assessing safety levels. On-board sensors, including eye-tracking cameras were employed to analyze the driver's visual distractions by extracting features through image analysis techniques.

The major contribution of this research is the consolidation of driver and driving state features that are inherent to the physiological and cognitive attributes of individual drivers and unique driving contexts that incorporate vehicle and traffic features. The key research question is: How does haptic attentional feedback affect the driver's reaction time, standalone as well as in combination with visual and auditory feedback? Is one modality as standalone or in combination the best way of providing attention warnings to the driver.

## 2. Method

### *2.1 Study Setting*

Autoliv's simulator was used to conduct similar tests to NHTSA (2011). The same modalities in the NHTSA's study (2011) were tested. Eye-tracking was used to study eye gazing before, during and after the warning.

### *2.2 Participants*

There was a total of 80 participants, 10 (between-group) participants in each of the eight modality levels. The participants had no prior knowledge of the warnings, the meaning behind the feedbacks or an animal that animated out of the forest. The ages were 19-65 with a goal of having an equal number of men and women. They approved of the ethical factors in relation to their participation.

### *2.3 Experiment conditions*

The eight experimental conditions were: Haptic (H), Visual (V), Auditory (A), Haptic-Visual (HV), Haptic-Auditory (HA), Visual-Auditory (VA), Haptic-Visual-Auditory (HVA), and No Feedback (NF).

The scenario (Fig. 1) was set up in a virtual remake of the AstaZero test-track using the Autoliv simulator. Participants were told to drive around the rural road of AstaZero at 70 km/h and keep right. Approximately three minutes in, the participant approached the warning-stretch, where they were about to get one of different warning combinations.

Haptic warnings were activated using a manual button, while visual and auditory were automatically activated using scripts. Therefore, visual queries were placed [Q] 50 m before the automatic activation of visual and auditory warnings [A] to accurately time the manual button click with the automatic warning. When passing the blue column [A] the hidden animal [B] started animating towards the road [C] (Fig. 2), while simultaneously the participant is being presented with one of four warning-combinations. While the test was aimed to study the effects of modalities and reaction times made, crash-statistics were noted


**Fig. 1.** The traffic area with a maximum of 70 km/h. *[Q]* warning zoon, *[A]* warning activation, *[B]* object

The scenario object (deer) had to be moved due to some technical irregularities. This seemed to affect the result (later slide) and therefor the VA group was excluded from the data analysis. VA will be included in the graphs but cannot be taken into consideration as the eye-tracking results show that this change had an effect.

Reaction times were measured from the time of warning to the participant braking -using automated data to extract timestamps and confirming them using eye-tracking. The participants had no prior knowledge of the warnings, the meaning behind the feedbacks or animal that animated out of the forest.

During the drive, participants saw oncoming traffic, parked vehicles and people on the side of the roads to avoid sudden suspicion to the visual queries.

Eye-tracking glasses were used to observe eye gazing and evaluate possible results.

**Fig. 2.** The traffic area with a maximum of 70 km/h. *[Q]* warning zoon, *[A]* warning activation, *[B]* object

During the drive, participants saw oncoming traffic, parked vehicles and people on the side of the roads to avoid sudden suspicion to the visual queries. While the purpose of the tests was aimed to study the effects of attention and inattention modalities according to reaction times, crash statistics were also recorded.

## 3. Results

The primary results are the reaction times (Fig. 3) and crash percentages (Fig. 4) in relation to the different perceptual warnings.

### *3.1 Reaction time (RT) and crash percentage*

The fastest reaction time was VA (1.12) followed by H (1.37). But due to the technical irregularities VA was excluded from any kind of statistical test. The slowest reaction time was V. The main effect was not significant, one-way ANOVA main effect, $F(7,72) = 1{,}98$, $p = 0.135$, $n^2 = 0.142$. There is, however, a large difference between the haptic and the haptic-visual-auditory conditions in Fig. 3, which was statistically tested as well. The RT difference between those two conditions was significant, via the multiple comparison post-hoc test LSD, $p(0{,}026)$. There was also a similar significant difference between H and NF, $p(0{,}036)$ and H and V $p(0{,}028)$.

The crash/non-crash percentages for each condition show a large difference between NF and H. For NF, crashes were 90%, and for H, 40%. And H was the only condition where there were fewer crashes than non-crashes.

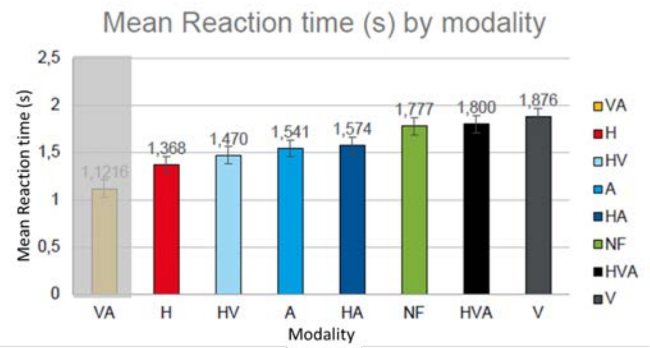



**Fig. 3.** Mean reaction times for the eight perceptual warning conditions.

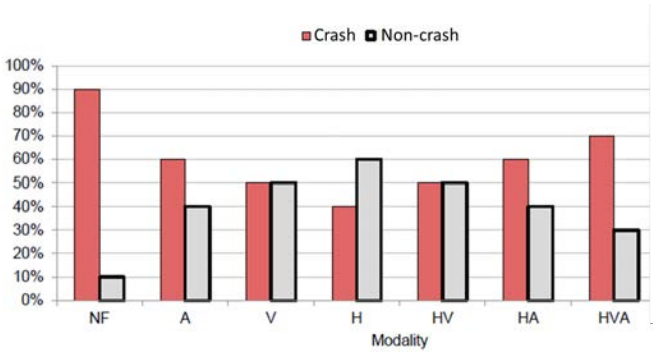



**Fig. 4.** Mean percentages for crash/non-crash for the eight perceptual warning conditions.

## 4. Discussion and Conclusions

The visual and auditory information was provided via the criteria in ITS Informal Group (2011). The other conditions that were significantly slower than the haptic factor were the three different combinations with the haptic factor. Given these results, haptic warning information has a positive effect for drivers' responce and reduced crashes. The other condition very near haptic information is the auditory condition for both reaction time and crash percentage. The limitation of this study is the one traffic situation that was investigated when a deer can come out on the road. A key question and further development then has to do with other traffic situations and driver interaction with perceptual attention factors to reduce driver distraction. This will be further studied.

Through the application of machine learning algorithms, future ADAS systems in collaboration with HMI recognize various driving patterns that were hidden from previous traffic safety solutions. This includes the combination of driver distraction levels, as well as vehicle dynamics and ambient traffic to enable the derivation of more precise safety indicators. These indicators are then utilized for HMI functions.


## 5. Acknowledgments

Two projects have contributed to this study: Pre-study (2022 01652) FFI-Vinnova - Turning Weak Evidence into Strong Safety Performance Scores (Tweaks); Vinnova (2023-03056) CogniDrive: Support Cognitive Abilities for Predictive Processing through Driver Monitoring. Autoliv AB is the study producer.


**References**


Casner, S.M., Hutchins, E.L., Norman, D. (2016). The challenges of partially automated driving. *Association for Computing Machinery*. 59(5): 70-77.

Favaro, F.M., Eurich, S.O., Rizvi, S.S. (2019). "Human" Problems in Semi-Autonomous Vehicles: Understanding Drivers' Reactions to Off-Nominal Scenarios. *International Journal of Human–Computer Interaction*. 35: 956-971.

Forkenbrock, G., Snyder, A., Heitz, M., Hoover, R., O'Harra, B., Smith, L. (2011). A Test Track Protocol for Assessing Forward Collision Warning Driver-Vehicle Interface Effectiveness (DOT HS 811 501). *National Highway Traffic Safety Administration*.

Gaffary, Y., Lécuye, A. (2018). The Use of Haptic and Tactile Information in the Car to Improve Driving Safety: A Review of Current Technologies. *Human-Media Interaction*. 5(5).

ITS Informal Group (2011). Guidelines on establishing requirements for high-priority warning signals with GRE comments to WP.29-150-22. *United Nations Economic Commission for Europe*.

Michelaraki, E., Katrakazas, C., Kaiser, S., Brijs, T., & Yannis, G. (2023). Real-time monitoring of driver distraction: state-of-the-art and future insights. *Accident Analysis & Prevention*, 192, 107241.

Masello, L., Sheehan, B., Castignani, G., Shannon, D., & Murphy, F. (2023). On the impact of advanced driver assistance systems on driving distraction and risky behaviour: An empirical analysis of irish commercial drivers. *Accident Analysis & Prevention*, 183, 106969.

Riener, A., Jeon, M., Alvarez, I., & Frison, A. K. (2017). Driver in the loop: Best practices in automotive sensing and feedback mechanisms. *Automotive user interfaces: creating interactive experiences in the car*, 295-323.

Xin, X., Wang, Y., Liu, N., Yang, W., Dong, H., Liu, W. (2021). Research on In-Vehicle Haptic Interactions as Crucial Resources for Driver Perceptions. *23rd HCI International Conference*. 373–388.

# Assessing the cognitive demand of In-Vehicle Infotainment System functions

Isabelle Roche Cerasi [1*], Dagfinn Moe [1], Ann-Helen Hansen [2] Bård Morten Johansen [2]

[1] SINTEF Community, Department of Mobility, S.P. Andersens vei 3, Trondheim, Norway.
[2] Norwegian Council for Road Safety, Trygg Trafikk, Tullinsgate 2, Oslo, Norway.
[*]isabelle.roche-cerasi@sintef.no

**Abstract: The objective of this study was to understand the cognitive demand in using In-Vehicle Infotainment system functions while driving. A total of 40 students of the Driving instructor program at the Nord University in Norway were asked to drive a car with double pedal set and a safety instructor on a specific route. They voluntarily performed four tasks on the touchscreen of different difficulty levels and the rides were registered with the use of eye-tracking glasses. The four tasks were: (1) to choose a radio channel, (2) to use the air conditioning system and fix temperatures for the driver and the passenger (3) to use the navigation system and (4) to listen to music on a streaming platform. The preliminary results suggest that the IVIS functions while driving require many switching sequences with the navigation system being the more demanding task. The relation between the speed distribution and the attention directed to the screen while driving provide an understanding of the risk incurred by the drivers. The cognitive load associated with switching sequences increases the risk of being surprised and affect the ability to react to risky situations. An adverse effect is the reduction in the "prediction error" time the brain needs to avoid surprises in traffic and to correct the plan and the execution of the actions. This paper contributes to better knowledge about the issues related to the use of car touchscreens with the help of eye-tracking and will be of high interest for the road safety community.**

## 1. Introduction

Our brain builds cognitive maps to make it possible for us to know where we physically are in a landscape ("where am I?"), orientate ourselves and move to a location ("where am I going?"). These mental representations of our spatial environment are connected to our place and grid cells in our brain hippocampus area and entorhinal cortex providing our own navigation system (Moser and Moser, 2017).

When significant changes occur in our environment, our cognitive maps are continuously updated and replaced with new ones. Such changes are called "cognitive remapping" in neuroscience research studies (Latuske, 2018; Green, 2022; Sugars, 2019). We, humans continuously predict and re-plan, asking ourselves internally over and over, what is going to happen? where? when and how? (O`Keefe and Nadel, 1978; Buzsáki and Moser, 2013; Nadel, 2021). Performing the driving tasks consists of cognitive remapping sequences and the same applies when drivers use In-Vehicle Infotainment Systems (IVIS), they must find their way through the menu and its options. Using such a system while driving requires switching back and forth between the traffic situation and the touchscreen and to repeatedly process cognitive remapping, updating our working memory and long-term memory.

Previous studies have investigated the driver-interaction with IVIS systems or mobile phones. The findings showed that interacting with systems requires high levels of both cognitive workload and interaction times with the system (Biondi et al., 2019, Buschholtz et al., 2023). In driving simulator studies, results demonstrated that drivers adapt their behaviour and move back their attention to the road when they perceive that the situations required it (Platten et al., 2013). According to the European Transport Safety Council (ETSC, 2023), long glances (greater than 2.0s) away from the road are correlated with increased accident risk. The National Highway Traffic Safety Administration published in 2013 the same guidelines for evaluating eye gaze behaviour while driving (NHTSA, 2013), indicating that 85% of eye glance durations away from the roadway should be 2s or less. The occlusion testing theory implies that the driver should also be able to complete a task with a Total Eyes-Off-Road Time of 12s (ISO 16673). Recent studies concluded that more field tests are needed to contribute to better HMI design and to understand the effects of modern touchscreens on road safety (Tinga et al., 2023, Buchholtz et al., 2023).

The objective of this study was to understand the visual, manual and cognitive workload of the drivers while using different IVIS-touchscreen functions. The use of an eye-tracking system in real traffic conditions allows us to evaluate their switching eye movement patterns and to relate these results to their speed choice.

## 2. Method

### *2.1 Participants*

A sample of 40 voluntary students at the Nord University in Stjørdal (Norway) participated in the study. They all followed the 2 year-License B Driving Instructor Education program to become instructors of driving schools.

### *2.2 Procedure*

The participants were asked to perform four tasks while driving on a specific route: (1) to choose a radio channel, (2) to use the air conditioning system and set temperatures for the driver and the passenger (3) to use the navigation system and (4) to listen to music on the streaming platform. They had the choice to perform the tasks or not.

The schedule times on the route for performing the tasks were defined in advance considering the location, the traffic situation and the use of the touchscreen functions.

The drivers had to wear eye tracking glasses and drive an automatic car with two sets of pedals (also used in the context of their education program). The role of the safety instructor on the passenger seat was to provide the task instructions at the right time and ensure safety. The speed of the car was registered from the instrument panel.

After the driving, the participants had to fill a form based on the 7 point-scale NASA RTLX assessment tool to give their opinion about using the IVIS-touchscreen functions while driving.

### *2.3 Materials*

The eye tracking system was provided by Tobii AB. The system consists of the Tobii Pro Glasses 2 with the Tobii Pro Lab (Version 1.162) [Computer software], Danderyd, Sweden: Tobii AB. The system collects raw eye movement data points, and each data point is identified by a timestamp and coordinates. The eye movements alternate between saccades (20-40 ms) and fixation points (100-600 ms).

The touchscreen of the car used for the study is a 10-inches HD touchscreen with 7 touch function buttons. A phone was connected to the screen via Bluetooth to use a streaming music platform.

## 3. Results

### *3.1 Eye tracker validation*

The validation is a procedure to assess the eye tracking data quality by calculating the eye-tracking samples. The system has a 50 Hz sampling frequency that generates 50 samples per second. Ideally if the system uses all the samples to calculate the gaze points, the gaze sample would be 100%. However, blinking often causes 5-10% data loss during a recording (Tobii AB, 2021).

The gaze samples for each driver when performing a task are calculated as follows: Gaze sample = (Number of unclassified samples + saccades + fixations)/Theoretical gaze samples and gaze samples with poor data quality (under 90%) are discarded from the study.

### *3.2 Eye tracker data*

The preliminary results show that the task that required the most driver attention was to enter an address in the navigation system.

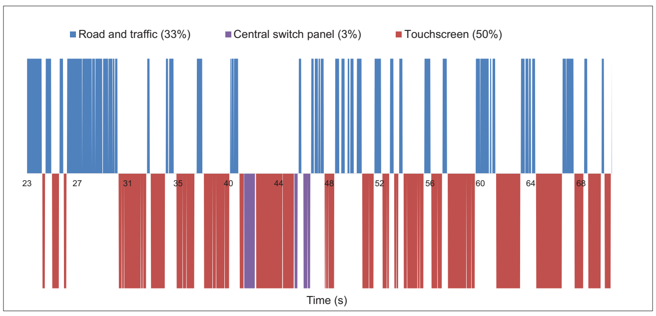



***Fig. 1.*** *Example of the attentional distribution over time (fixation points) of a participant while using the navigation system.*

Figure 1 presents an example of the attentional distribution of a participant when performing the task in an urban setting. The driver focused 50.4% of his attention on the screen and 33.3% on the road and traffic (including mirrors). The remaining time left is the fixation points on the central switch panel (3.4%), the saccades or the missing points due to eye blinking (12.9%).

Statistics will be performed for the whole sample to investigate the attentional distribution for the four tasks to evaluate the demand and the effort required of the drivers and the inherent accidental risk associated to their speed adjustments while using the different touchscreen functions.

### *3.3 NASA RTLX assessment tool*

Figure 2 shows that the mean perceived stress scores are approximately 4 for 5 of the indicators, except for the temporal demand (M=2.86, SD=1.72) and frustration level (M=3.14, SD=1.89). The results are spread out from the average score indicating differences in perceived demand and risk perception among the drivers.

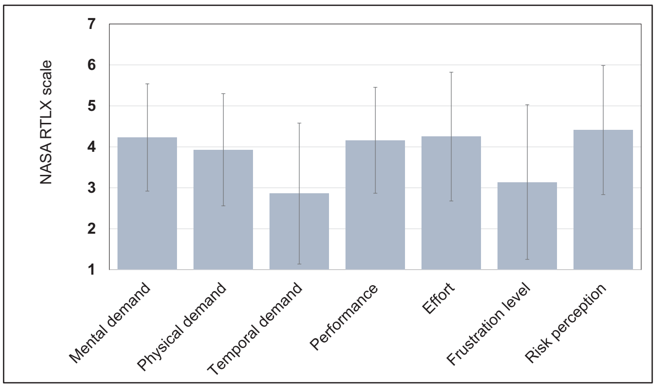



***Fig. 2.*** *NASA RTLX results for all the drivers (n=40).*

Figure 3 indicates that the tasks performed by the drivers are perceived with different levels of cognitive difficulty. Activating wipers is found to be easy whereas entering an address is evaluated as the most demanding task in terms of attentional load. The results are spread out indicating differences in perceived cognitive abilities among the drivers.

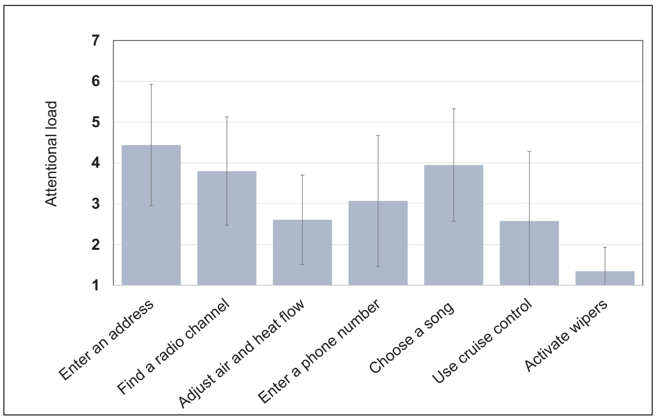



***Fig. 3.*** *Perceived attentional loads for all the drivers while performing the different tasks (n=40).*

## 4. Discussion

In this study, the tasks performed with the IVIS functions require many switching actions between the road and the screen. The cognitive load associated with back and forth switching sequences increases the risk for the drivers of being surprised and affect their ability to react to risky situations. An adverse effect is the reduction in the "prediction error" time the brain needs to avoid surprises in traffic and to correct the plan and the execution of the actions. The cognitive load associated with the IVIS tasks suggest a reduction in driver skills with drivers less likely to anticipate hazardous situations and adjust their behaviour. For example, drivers may become less precise, e.g. less speed control, less mirror control, and limited head movement.

## 5. Conclusions

The preliminary results suggest that the IVIS functions while driving require many switching sequences with the navigation system being the more demanding task. Please note that the final results will be provided in the oral presentation and the published journal paper. The relation between the speed distribution and the attention directed to the screen while driving provides an understanding of the risk incurred by the drivers.

## 6. Acknowledgments

This study was financed by the Norwegian Council for Road Safety, Trygg trafikk. Trygg trafikk and SINTEF would like to thank all the participants of the study and the professors at the Nord University for their collaboration in this study, Petter Helmersen Bogfjellmo and Kåre Robertsen. Additional thank to Abrar Hazoor, research scientist for his technical support and to An-Magritt Kummeneje, associate professor for her expertise and organisational skills.

## References

Biondi, F.N., Getty, D., Cooper, J.M., & Strayer, D.L. (2019). Examining the effect of infotainment auditory-vocal systems' design components on workload and usability. *Transportation Research Part F: Traffic Psychology and Behaviour, 62, 520-528.* https://doi.org/10.1016/j.trf.2019.02.006.

Buchholz, M., Wögerbauer, E., & Brandenburg, S. (2023). Comparison of the Box task combined with a tactile Detection Response Task and the Occlusion Method as measures of secondary task demand. *Transp. Research Part F.: Traffic Psychology and Behaviour, 97, 328-346.* https://doi.org/10.1016/j.trf.2023.07.019

Buzsáki, G., & Moser, E. (2013). Memory, navigation and theta rhythm, in the hippocampal-entorhinal system. *Nature Neuroscience, 16(2), 130-8.* DOI: 10.1038/nn.3304.

European Transport Safety Council (2023). Mandatory distraction warning systems won't detect most important types of distraction. Mandatory distraction warning systems won't detect most important types of distraction – ETSC

Green, L., Tingley, D., Rinzel, J., & Buzsáki, G. (2022). Action-Driving remapping of hippocampal populations in jumping rats. *PNAS, 119(26).* https://doi.org/10.1073/pnas.2122141119

International Standard Organisation. Road vehicles – Ergonomic aspects pf transport information and control systems – Occlusion method to assess visual demand due to the use of in-vehicle systems. ISO 16673:2017.

Latuske, P., Kornienko, O., Kohler, L., & Allen, K. (2018). Hippocampal Remapping and Its Entorhinal Origin. Frontiers. *Behavioural Neuroscience, 11.* https://doi.org/10.3389/fnbeh.2017.00253

Ma, J., Gong, Z., Tan, J., Zhang, Q., & Zuo, Y. (2020). Assessing the driving distraction effect of vehicle HMI displays using data mining techniques. *Transp. Research Part F.: Traffic Psychology and Behaviour, 69, 235-250.*

Moser, M.B., & Moser, E. (2017): *Where am I? Where am I going?* Scientific American. Special edition. DOI: 10.1038/scientificamerican0116-26

Nadel, L. (2021): The hippocampal formation and action at a distance. *PNAS, 118(51).* DOI: 10.1073/pnas.2119670118.

National Highway Traffic Safety (2013). Driver guidelines for In-Vehicle Electronic Devices. *Federal register. 18(81), 24818-24890.*

O'Keefe, J. & Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford Univ. Press, New York.

Platten, F., Milicic, N., Schwalm, M., & Krems, J. (2013). Using an infotainment system while driving – A continuous analysis of behaviour adaptations. *Transp. Research Part F.: Traffic Psychology and Behaviour, 21, 103-112.* https://doi.org/10.1016/j.trf.2013.09.012

Sugar, J., & Moser, M.B. (2019): Episodic memory: Neuronal codes for what, where, and when. *Hippocampus.* 29,1190–1205. DOI: 10.1002/hipo.23132

Tinga, A.M., van Zeumeren, I.M., Christoph, M., van Grondelle, E., Cleij, D., Aldea, A., & van Nes, N. (2023). Development and evaluation of a human machine interface to support mode awareness in different automated driving modes. *Transp. Research Part F.: Traffic Psychology and Behaviour, 92, 238-254.* https://doi.org/10.1016/j.trf.2022.10.023

Tobii AB (2021). Tobii Pro Lab User Manual (Version 1.162). Tobii AB, Danderyd, Sweden

# A Comparison of Drowsiness Measures in Truck and Bus Naturalistic Driving Data: Streamlining Drowsiness Reduction

Susan Soccolich [1*], Rebecca Hammond [1], Matthew Camden [1], and Stuart Walker [1]

[1] Virginia Tech Transportation Institute, Virginia Tech, 3500 Transportation Research Plaza, Blacksburg, USA
[*]ssoccolich@vtti.vt.edu

**Abstract: Naturalistic driving data allows researchers an opportunity to better understand drowsy driving, through review of driver-facing video capturing driver behavior and eye closures as drowsiness occurs in the real world. The current study explored how different drowsiness measures impact fatigue determination (whether fatigue was observed) in driving samples and study estimates of fatigue prevalence, risk, and secondary task association for truck and motorcoach drivers. The data was previously collected and reduced in two completed naturalistic driving studies. Analyses investigated PERCLOS scores using 1 minute of data (PERCLOS 1) versus 3 minutes of data (PERCLOS 3). The study found the sample size of events with valid PERCLOS data increased by 8.94% when PERCLOS 1 criteria were used. Matching fatigue determination in PERCLOS 3 and PERCLOS 1 scores was found for 95.89% to 99.48% of truck and motorcoach driving samples. The risk of safety-critical event involvement when driving while fatigued was consistent for truck drivers when using PERCLOS 1 or PERCLOS 3 to determine fatigue. However, for motorcoach drivers, SCE risk estimates were affected by PERCLOS measure used. The study investigated methods to lessen the effort of fatigue data reduction in future studies and obtain the most valuable dataset at the lowest time and cost budget. The study found that, a targeted fatigue reduction approach that includes ORD for all events and targeted PERCLOS 3 or PERCLOS 1 reduction for events that meet or exceed an ORD threshold can reduce the cost of fatigue reduction while maintaining the advantage of ORD reduction.**

## 1. Introduction

Every year, drowsy and fatigued driving contributes to thousands of crashes and their resulting injuries and fatalities (National Center for Statistics and Analysis, 2017). Drowsy driving prevalence in truck-involved crashes has been estimated at 13% (Federal Motor Carrier Safety Administration, 2005). However, these estimates likely underestimate how often drowsiness contributes to crashes.

Naturalistic driving data allows researchers an opportunity to better understand drowsy driving, through review of driver-facing video capturing the driver's behavior and eyes as it occurs in the real world. In a recently completed naturalistic driving study, reduction of a subset of events included two drowsiness measures successfully used in naturalistic driving data: Observer Rating of Drowsiness (ORD) (Wiegand, McClafferty, McDonald, & Hanowski, 2009) and manual percentage of eye closure (PERCLOS) (Wierwille & Ellsworth, 1994).

The reduction of both ORD and PERCLOS 3 in Hammond et al. (2021) provided a unique opportunity to compare drowsiness measures for strengths, weaknesses, similarities, and differences. The current study explored how drowsiness measures impact study estimates of fatigue prevalence and risk for truck and motorcoach drivers. The study examined the relationship across ORD, PERCLOS 1, and PERCLOS 3 drowsiness measures to identify ways to obtain the most valuable dataset at the lowest cost to time and budget.

## 2. Method

### *2.1 Data Source*

The current study utilized naturalistic driving data collected during the Onboard Monitoring System Field Operational Test (OBMS FOT) study (Boyle, Guo, Hammond, Hanowski, & Soccolich, 2016) and further reduced during the Naturalistic Driving Study (Hammond et al., 2021). The study included 172 truck driver and 73 motorcoach driver participants. Safety-critical events (SCEs) and baseline driving epochs (BLs) underwent reduction for secondary task engagement, driver drowsiness, and more.

### *2.2 Drowsiness Measures*

Driver drowsiness reduction performed in Hammond et al. (2021) included ORD scored over one minute of data and manual PERCLOS scored over three minutes of data (PERCLOS 3). ORD measures drowsy driving with a subjective assessment of the driver (Wiegand, McClafferty, McDonald, & Hanowski, 2009) and is rated on a scale of 0 to 100. Manual PERCLOS is the percentage of time when the driver's eyes are "80 to 100 percent closed" (Wierwille & Ellsworth, 1994) and is coded by reviewing each sync for 1 minute (PERCLOS 1) up to 3 minutes (PERCLOS 3) leading up to an SCE/BL. ORD and PERCLOS scores can be used to make a "fatigue determination" about a moment of driving-when the ORD score or PERCLOS score reaches or exceeds a specified value, the driver is said to be fatigued (and conversely, if the scores are under the specified value the driver is said to not be fatigued).

In the current study, PERCLOS 1 scores were calculated from PERCLOS 3 data for all SCEs/BLs. That is, the PERCLOS 1 values used in this report were calculated from a one-minute subset of the original PERCLOS 3 coded data. The fatigue determination was made for all SCEs/BLs for each of the drowsiness measures calculated (ORD, PERCLOS 3, and PERCLOS 1).

### *2.3 Analysis Approach*

Hammond et al. (2021) study assessments of the relationships between SCE involvement, driver distraction, and drowsy driving (using PERCLOS 3) were reassessed using PERCLOS 1. Generalized linear mixed-effect models and resulting odds ratios (ORs) and 95% confidence intervals (CIs) were used to estimate the risk of SCE involvement during fatigued driving.

The final analysis investigated how a multifaceted drowsiness reduction approach could be used to maximize drowsiness data collection and reduce drowsiness reduction costs. ORD ratings for PERCLOS 3 and 1 fatigued events and non-fatigued events were compared and the ORD rating threshold that the PERCLOS fatigued events met or exceeded was identified. Estimates of reduction time and cost were calculated and compared for several drowsiness reduction approaches, including performing ORD on all events and targeted PERCLOS reduction on events meeting or exceeding the ORD rating threshold (ORD and targeted PERCLOS 3, ORD and targeted PERCLOS 1).

## 3. Results

### *3.1 Investigating Prevalence and Risk of Fatigue using PERCLOS 1*

When using PERCLOS 1, the prevalence of fatigue in BL data was found to be 0.52% for motorcoach data (11 BLs) and 3.95% for truck data (85 BLs). The prevalence of fatigue in SCE data was found to be 1.42% for motorcoach data (115 SCEs) and 12.69% for truck data (182 SCEs). Driving while fatigued was found to increase risk of SCE involvement by 2.31 times compared to driving without fatigue for truck drivers [95% CI = (1.69, 3.15)]- a finding consistent with the PERCLOS 3 finding in Hammond et al. (2021). Motorcoach drivers showed no significant change in SCE risk when driving while fatigued compared to driving without fatigue [OR = 2.12, 95% CI = (0.94, 4.78)]. This result was not consistent with the Hammond et al. (2021) finding based on PERCLOS 3 data.

### *3.2 Identifying Fatigue Reduction Options and Evaluating Reduction Time and Cost*

The first step in identifying an ORD rating score interval for targeted PERCLOS reduction was to understand the distribution of ORD ratings for events marked as fatigued using PERCLOS 3 and PERCLOS 1 score calculations. All but one PERCLOS fatigued event with ORD ratings of “slightly drowsy” had ORD rating scores of 27 or higher. If an ORD rating score of at least 27 was set as a cutoff for performing PERCLOS reduction, 58.38% events in the current study would receive PERCLOS reduction. If the ORD rating score cutoff was lowered to at least 25 to include a small buffer, 61.90% of events in the current study would receive PERCLOS reduction.

Based on current reduction rates, the cost for drowsiness reduction for the full event dataset ($n$ = 6,763) would be approximately $66,953.70 for ORD reduction, $89,271.60 for PERCLOS 3 reduction, and $31,433.66 for PERCLOS 1 reduction. Table 1 compares the reduction costs for ORD and targeted PERCLOS options to the approach of ORD and PERCLOS for all events. A targeted approach using PERCLOS 1 (the most cost-efficient option) is 55.31% of the cost of all events receiving ORD and PERCLOS 3 reduction (the highest cost option).

***Table 1*** *Proportion of reduction cost for ORD and targeted PERCLOS option compared to all ORD and PERCLOS option.*

| Proposed Fatigue Reduction Option | Comparison Option | Proportion of Cost |
|---|---|---|
| All ORD + Targeted PERCLOS 3 at ORD Threshold 25 | All ORD + All PERCLOS 3 | 78.23% |
| All ORD + Targeted PERCLOS 1 at ORD Threshold 25 | All ORD + All PERCLOS 1 | 87.83% |
| All ORD + Targeted PERCLOS 1 at ORD Threshold 25 | All ORD + All PERCLOS 3 | 55.31% |

## 4. Conclusions

Drowsy and fatigued driving is a critical problem, resulting in thousands of crashes and fatalities every year. Naturalistic driving research provides critical insight on the characteristics of drowsy driving and the frequency and severity with which it occurs. The use of PERCLOS 1 can increase the available event sample size for analysis.

Fatigue prevalence in BLs was slightly different based on the PERCLOS method. Fatigue prevalence as measured using PERCLOS 1 was higher in motorcoach and truck SCEs compared to BLs. A similar range of fatigue prevalence has been observed in other naturalistic driving studies (Hammond et al., 2016; Klauer et al., 2006; Dingus et al., 2016; Owens et al., 2018). Truck drivers showed statistically significant increased risk of SCE involvement when fatigued, regardless of PERCLOS method used to measure fatigue. Increased risk of safety event involvement when fatigued has been found in several driving studies (Hanowski, 2000; FMCSA, 2005; Dingus et al., 2006; Dingus et al., 2016).

The assessment of drowsiness measurement methods for reduction cost and time identified an opportunity to limit costs and obtain full, rich datasets using a targeted reduction approach. Performing ORD for all events and PERCLOS 1 for targeted events cut reduction costs in half compared to all events receiving ORD and PERCLOS 3 reduction. A targeted approach benefits from the strengths of multiple drowsiness measurement methods. Researchers evaluating the best drowsiness measurement method for their study will need to consider how the different methods will impact range of drowsiness captured, number of events to assess, and time and monetary budget for fatigue reduction. Researchers may want to consider for what purpose they are assessing drowsiness and whether it is important to capture early stages of drowsiness and fatigue-fighting or fatigue-managing behaviors or to capture fatigue in more advanced stages.

## 5. Acknowledgments

The authors of this report would like to acknowledge the support of the stakeholders of the National Surface Transportation Safety Center for Excellence (NSTSCE): Zac Doerzaph from the Virginia Tech Transportation Institute; John Capp from General Motors Corporation; Terri Hallquist from the Federal Motor Carrier Safety Administration; Mike Fontaine from the Virginia Department of Transportation and the Virginia Transportation Research Council; and Melissa Miles from State Farm Insurance. The authors of this report would like to acknowledge the support of the Division for Freight, Transit, and Heavy Vehicle Safety at the Virginia Tech Transportation Institute in funding this research. The authors would like to thank Dr. Richard J. Hanowski for contributing to this research study, through guidance, support, and review.

The NSTSCE stakeholders have jointly funded this research for the purpose of developing and disseminating advanced transportation safety techniques and innovations.

**References**

Boyle, L. N., Guo, E. H., Hammond, R. L., Hanowski, R. J., & Soccolich, S. A. (2016). *Performance assessment of an onboard monitoring system for commercial motor vehicle drivers: A field operational test (Report No. FMCSA-RRR-15-019).* Washington, DC: Federal Motor Carrier and Safety Administration, USDOT.

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., 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, *113*(10), 2636-41.

Federal Motor Carrier Safety Administration. (2005*). Report to Congress on the Large Truck Crash Causation Study.* Washington, D.C: U.S. DOT. https://rosap.ntl.bts.gov/view/dot/61146

Federal Motor Carrier Safety Administration. (2015). *CMV driving tips- driver fatigue.* https://www.fmcsa.dot.gov/safety/driver-safety/cmv-driving-tips-driver-fatigue

Hammond, R. H., Soccolich, S. S., Han, S., Guo, F., Glenn, T. L., & Hanowski, R. J. (2021). *Analysis of naturalistic driving data to assess distraction and drowsiness in drivers of commercial motor trucks and buses*. Washington, DC: Federal Motor Carrier Safety Association: Washington, USDOT. https://rosap.ntl.bts.gov/view/dot/57153

Hammond, R. L., Hanowski, R. J., Miller, A. M., Soccolich, S. A., & Farrell, L. J. (2016). *Distraction and drowsiness in motorcoach drivers (Report No. FMCSA-RRR-15-017).* Washington, DC: Federal Motor Carrier Safety Administration, USDOT.

Hanowski, R. J. (2000). *The impact of local/short haul operations on driver fatigue* (dissertation). Virginia Tech.

Hanowski, R. J., Blanco, M., Nakata, A., Hickman, J. S., Schaudt, W. A., Fumero, M. C., Olson, R. L., Jermeland, J., Greening, M., Holbrook, G. T., Knipling, R. R., & Madison, P. (2008a). *The drowsy driver warning system field operational test: Data collection methods* (Report No. DOT HS 811 035). Washington, DC: National Highway Traffic Safety Administration, USDOT.

Hanowski, R. J., Bowman, D., Alden, A., Wierwille, W., & Carroll, R. (2008b). *PERCLOS+: Development of a robust field measure of driver drowsiness.* New York, NY: 15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual Meeting.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data* (Report No. FHWA-HRT-04-138). Washington, DC: National Highway Traffic Safety Administration, USDOT. http://www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/Crash%20Avoidance/2006/DriverInattention.pdf

National Center for Statistics and Analysis. (2017, October). *Drowsy driving 2015 (Crash Stats Brief Statistical Summary* (Report No. DOT HS 812 446). Washington, DC: National Highway Traffic Safety Administration, USDOT. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812446

National Highway Traffic Safety Administration. (n.d.) *Drowsy driving*. https://www.nhtsa.gov/risky-driving/drowsy-driving

National Institute for Occupational Safety and Health. (2023). *Work and fatigue.* Centers for Disease Control and Prevention. https://www.cdc.gov/niosh/topics/fatigue/default.html

National Surface Transportation Safety Center for Excellence. (2017). *CMV driving safety: Driver drowsiness and fatigue.* Blacksburg, VA.

Owens, J. M., Dingus, T. A., Guo, F., Fang, Y., Perez, M., McClafferty, J., & Tefft, B. (2018, February). *Prevalence of drowsy driving crashes: Estimates from a large-scale naturalistic driving study.* (Research Brief.) Washington, DC: AAA Foundation for Traffic Safety. Retrieved from: http://aaafoundation.org/wp-content/uploads/2018/02/FINAL_AAAFTS-Drowsy-Driving-Research-Brief-1.pdf

Wiegand, D. M., McClafferty, J. M., McDonald, S. E., & Hanowski, R. J. (2009). *Development and evaluation of a naturalistic observer rating of drowsiness protocol.* Blacksburg, VA: The National Surface Transportation Safety Center for Excellence.

Wierwille, W. W. & Ellsworth, L. A. (1994a). Evaluation of driver drowsiness by trained observers. *Accident Analysis and Prevention*, *26*(5), 571-581.

Wierwille, W. W., Ellsworth, L. A., Wreggit, S. S., Fairbanks, R. J., Kirn, C. L. (1994b). *Research on vehicle-based driver status/performance monitoring: development, validation, and refinement of algorithms for detection of driver drowsiness* (Report No. DOT HS 808 247). Washington, DC: National Highway Traffic Safety Administration, USDOT.

Wierwille, W. W. (1999). *Historical perspective on slow eyelid closure: Whence PERCLOS? Ocular measures of driver alertness* (Report No. FHWA-MC-990136). Washington, DC: Federal Highway Administration, USDOT.

# Innovative Road Safety Education Program
# "Children, attention and cycling"

Isabelle Roche Cerasi [1], Dagfinn Moe[2*], Jan Petter Wigum[3], Bård Morten Johansen [4], Ann-Helen Hansen [5]

[1] SINTEF community, Dept of Mobility and Economics, Trondheim, Norway, *E-mail: isabelle.roche-cerasi@sintef.no*
[2] SINTEF community, Dept of Mobility and Economics, Trondheim, Norway, *E-mail: dagfinn.moe@sintef.no*
[3] Nord University, Stjordal Campus, Road Traffic Section, Norway, *Email: jan.p.wigum@nord.no*
[4] Trygg Trafikk, Tullinsgate 2, Oslo, Norway, *E-mail: johansen@tryggtrafikk.no*
[5] Trygg Trafikk, Tullinsgate 2, Oslo, Norway, *E-mail: ahhansen@tryggtrafikk.no*

**Abstract: The main goal of the project was to develop, test, and evaluate a new training model for bicycle traffic education based on the scientific concept Mind, Brain and Education (MBA), that utilizes the interaction between neuroscience, psychology, and pedagogy. Distribution of attention has an important impact on children´s behaviour in mixed traffic situations. Learning how to properly use attention in complex traffic situations has never been part of traditional road safety education in Norway. The Norwegian Council for road safety, Trygg Trafikk, with SINTEF and Nord University developed a new road safety education program based on the last findings in Neuro-Education. The training model is anchored in knowledge about the brain's development of cognitive maps, serving as the basis for navigation and episodic memory. The method developed for stimulating schoolchildren' reflection on traffic safety issues and three concepts: risk, orientation, and attention. SINTEF compared the new education program with the one currently in place in Norway at its Virtual Reality laboratory. The results showed that the experiment group who participated in the new education program orientated themselves and used their attention better than their counterparts in the control group who followed the traditional program.**

## 1. Introduction

The teaching model presented in the following has focused on the development of children's attention and risk perception. The Norwegian Council for road safety with SINTEF and Nord University developed a new road safety education program based on the last findings in Neuro-Education.

Traffic rules, signs and regulations have been an essential part of the school road safety education program over the past years. Proper use of attention in complex traffic situations has not earlier been a subject in this kind of training programs.

The new education program was compared with the one currently in place in Norway at SINTEF´s Virtual Reality laboratory. Our research shows that attention training can increase children´s ability to orient themselves and gain better self-regulation.

## 2. Method

The program was evaluated with two groups of 5th grade pupils; an experiment group exposed to the new training program, and the control group exposed to the standard program already in use. A bicycle simulator and a Head Mounted Display (HMD) with an integrated Tobii eye tracking system, were connected to the Virtual Reality.

Fig. 1: SINTEF´s Virtual Reality laboratory with eye tracking system

The virtual environment was an identical copy of the traffic center facility. The traffic center is a training facility for schools consisting of a miniature traffic system with intersections, traffic lights and signs. Children wearing glasses (issued with VR glasses) and children not used to bicycle to school were excluded from the study.

The HMD with eye tracking system allowed registering of time stamped fixation points, total fixation duration, and time to first fixation for each Area of Interest (AOI) in the virtual world. The AOIs have been chosen on the basis of risk assessments at road junctions and how this subject is trained in driver education.

A didactic cognitive teaching approach was chosen to develop schoolchildren's attention and ability to assess risks and construct cognitive maps (spatial representations of the environment). The approach required teachers to problematize complex traffic situations and control the instructional stimuli to engage children's reflection and actions. The teachers therefore had to undergo a guidance course in order to be able to carry out the teaching in this way.

As stated by Piaget (1973), knowledge emerges through active participation and curiosity. This interactive approach, or called constructivism explained that the cognitive development is formed based on interactions between brain cognitive functions and the environment (Inhelder and Piaget, 2013). Children are active learners and thinkers with problem-solving skills. The teacher is then a facilitator to boost children' reflective and critical thinking.

2.1 The new teaching model

The new model was elaborated by the research institute, SINTEF, The Norwegian Council for road safety and Nord University with municipal representatives (teachers) of road safety education. Prior to the activities, a guide was prepared for the teachers to fulfill their need for subject terminology and knowledge. They also had meetings with the involved parties and participated in the discussion for the elaboration of the model. These are the central activities of the model:

1. Start up: Schoolchildren become familiar with the concepts of attention, orientation, and risk assessment. They reflect on a video showing children cycling to school and taking risks at an intersection with reduced visibility.

2. Cycling exercises: They practice cycling skills such as keeping their balance and braking properly at the right place and time.

3. Cognitive maps: They are divided in groups, develop cognitive maps of the environment. Groups identify and evaluate risk factors for each location with the help of a map of the area.

4. Teachers' guidance: The groups present their risk assessments to other groups and the teacher. They also reflect on the consequences of their risk assessment on the orientation and attention.

5. Cycling and attention: They practice their newly developed "cognitive maps" through cycling and planning for action and navigation on the traffic center path.

6. Termination: They express what they have experienced and whether they can draw any learning from their experience.

Fig. 2. Educational activities at the Eberg traffic center

Children develop "cognitive maps", identify risk factors, work in groups and work with their teachers. The overlapping nature of the children´s cognitive systems involved in navigation, episodic and semantic memory, imagination, was stimulating for predicting changes in the environment and planning actions.

The teaching took place in an environment where the children divided into groups had to map the risk factors under the supervision of a teacher. This exercise stimulates the most important cognitive functions associated to sustained attention and spatial orientation and increase the ability to form episodic memories.

Previous studies have shown that the brain plasticity is activated through physical exercise, such as playing computer games or experiencing meditation (Sommerville, 2016; Graybiel and Scott, 2019). A crucial factor for effective learning in our new model was the use of short-term learning experiences with physical activity, curiosity, initiative, and entertainment.

## 3. Results

The results clearly showed that the experiment group who participated in the new education program had better orientation in the traffic environment and used their attention in a better way than the children in the control group who followed the traditional program. Fig. 5 below shows the proportion of observers of the nine AOIs per group. There are significant differences between the groups for 6 AOIs, excluding AOI 4, 6 and 7, and assuming that the cut-off point for the statistical results is ap-value of 0.05.

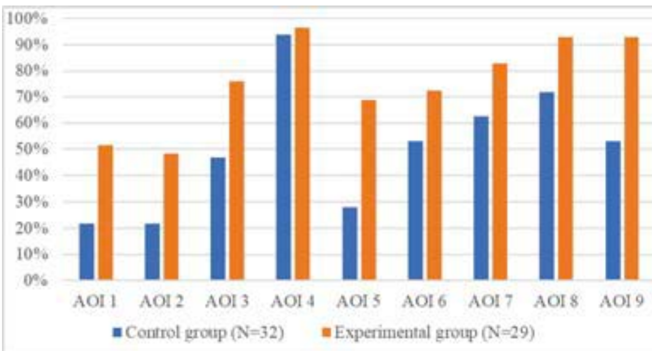


Fig. 3 Proportion of children in the groups who fixated the targeted AOIs.

The results showed a significant difference between the two groups, indicating that the new model had a higher impact on the executive functions of the children´s brain. The fixation time of the experimental group was found to be significantly longer in all intersections than for the control group, which may be an indication that children of the experimental group, had a more planned behaviour and systematically better brain collection of information. This also results in better speed adaption and breaking readiness.

Fig. 6 below shows the average fixation duration per group and per AOI. For all the AOIs (except AOI 8), a larger proportion of children in the experimental group fixated on average the areas longer than in the control group.

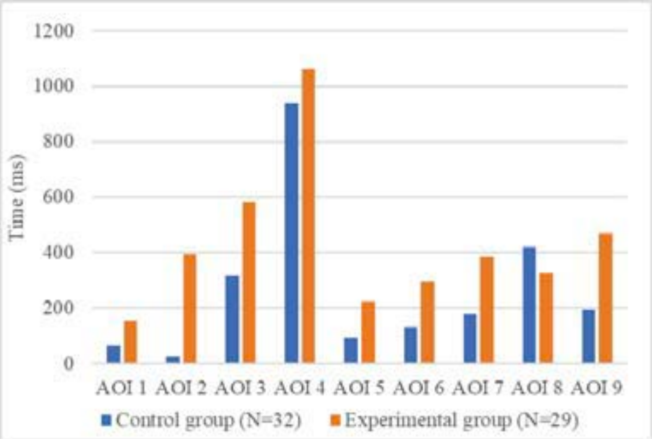



Fig.4 Average fixation duration (ms) for the groups.

## 4. Discussion

Previous studies showed that bicycle safety education increases children's knowledge and reduces the number of injuries and fatalities of cyclists (Hooshmand et al., 2014). Children's executive attention network is not fully developed, and this affects their ability to filter irrelevant information while cycling (Roche-Cerasi et al., 2017).

The new education program stimulated children to construct cognitive maps and to develop knowledge for assessing traffic risks. The virtual reality laboratory provided a safe environment to easily collect information about their performances without exposing the children to any dangerous traffic situations. After completing the traditional and new training programs, the effects were evaluated in the VR environment. The results showed a significant difference between the two groups, indicating that the new model had a higher impact on the executive functions of the children brain.

## 5. Conclusions

We can conclude that the new model was more effective in helping the children to focus their attention at the right time in the right places, to orientate themselves and to behave in a safer way when cycling. The present study showed that it is possible and necessary to change the traditional road safety education program by incorporating knowledge from neuroeducation.

## 6. Acknowledgments

The authors of this paper would like to thank The Norwegian Council for road safety and Ida Neergaard for their support, motivation, and commitment.

We also thank the teachers and the children who participated in this study, the County Council of Transport and Traffic Safety in Trøndelag, Nordland and Østfold for their financial support.

## References

Piaget, J. (1973). To understand is to invent: The future of education. (G.Roberts,Trans.) NY: Grossman Publishers.

Graybiel, A.M and T.G. Scott (2019). The Striatum. Where Skills and Habits Meets. Published by Cold Spring Harbor Laboratory Press.

Hooshmand J, G. Hotz, V. Neilson, L. Chandler (2014). Bikesafe: Evaluating a bicycle safety program for middle school aged children. Accident Analysis & Prevention, Vol 66, pp 182-186.

Inhelder B., & Piaget, J. (2013). The growth of logical thinking from childhood to adolescence. Book ed. Routledge, London. https://doi.org/10.4324/9781315009674

Roche-Cerasi I., D. Moe, J. Skjermo, R. Opland, J.P. Wigum (2017) Go/nogo testing and road safety education for young cyclists. Securité routière : état des lieux et initiatives dans le monde. Lharmattan ed

Sommerville, L.H (2016): Searching for Signatures for Brain Maturity. What are we searching for? Neuron NeuroView 10.059.

# Analyzing Autonomous Driving Misuse through Eye-Tracking and Driver Behavior

Mi Chang [1], Do Wook Kang [1], Woojin Kim[1], Eun Hye Jang[1], Daesub Yoon[1*]

[1] Mobility UX Research Section, Electronics and Telecommunications Research Institute, Daejeon 34129, Republic of Korea
[*]eyetracker@etri.re.kr

**Abstract: Despite the necessity for driver intervention in Level 2 autonomous driving, accidents often result from driver misuse of functionalities. Existing research primarily focuses on environmental factors and malfunctions, overlooking the correlation between driver tendencies and misuse-induced behaviors. This study aims to address this oversight by analyzing driver behavior patterns associated with misuse scenarios. To achieve this goal, we first developed a vehicle simulator based on the misuse scenario and conducted driver experiments with 77 participants. Subsequently, we analyzed driver misuse behavior factors from the experiments, identifying four key factors. Thirdly, we analyzed the pre- and post-accident behavioral characteristics of 25 drivers involved in accidents through eye-tracking data analysis. This study methodologically contributes to analyzing the driver misuse factors based on the analysis of driver behavior characteristics at the time of accidents**

## 1. Introduction

Autonomous driving technology denotes the capability of a vehicle to autonomously navigate, aiming to ensure driving safety for the driver. In the classification system established by the Society of Automotive Engineers (SAE), Level 2 autonomous vehicles necessitate driver intervention in certain situations to uphold driving safety.

Despite the requirement for driver intervention, accidents occur due to driver misuse of functionalities. For instance, accidents may result from drivers engaging in drowsy driving while the autonomous driving function is activated. Additionally, accidents may occur due to drivers failing to understand the system's limitations and thus being unable to switch to manual driving when necessary. Therefore, effective driver monitoring is crucial for safe manual driving transitions, and it is essential to identify driver behaviors before accidents occur. Additionally, pre-emptively identifying misuse scenarios and quantitatively analyzing driver behaviors in these scenarios are necessary.

However, existing research on autonomous driving primarily focuses on simplistic behaviors such as whether the driver's eyes are open or closed, or leans heavily towards improving malfunctioning autonomous driving functionalities. Furthermore, there is a tendency to solely concentrate on the immediate surroundings before car accidents, neglecting a thorough analysis of the relationship between driver tendencies and behaviors resulting from misuse of functionalities.

To overcome these limitations, this study aims to analyze driver behavior factors based on driver misuse scenarios and conduct driver behavior analysis using eye movement data before and after accidents. The study develops a vehicle simulator based on misuse scenarios of autonomous driving and contributes to grouping driver behavior factors and extracting specific factors. Furthermore, the study contributes to analyzing the driver behavior based on behavioral characteristics before and after accidents using driver eye-tracking data.

## 2. Materials and Methods

This study collects and analyzes data using vehicle simulator experiments and surveys, as shown in Figure 1. The process involves setting up a simulator environment, conducting experiments to gather participant data, and analyzing survey responses and vehicle accident data, with a focus on eye-tracking data.

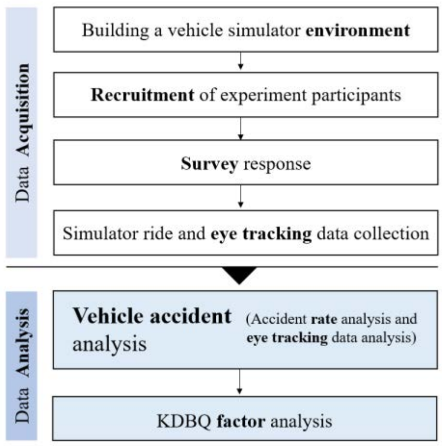



***Fig. 1.*** *Overview of the research process*

### *2.1 Experiment environment and scenario*


***Fig. 2.*** *Vehicle simulator environment*

To facilitate the experiment, a human driving cockpit simulator and human misuse scenario software were developed (Figure 2). The simulator, Hyundai Motors' New Grandeur, includes vehicle data protocols and interfaces. The Human misuse scenario, derived from previous research (Kim., et al, 2023), represents a situation in which drivers misuse Level 2 autonomous driving features due to misjudgement. It involves the system maintaining autonomy despite obstacles at speeds over 50 km/h, with drivers failing to intervene, assuming the vehicle will stop automatically. Data on driver gaze movements and vehicle control are collected at 20Hz throughout the scenario.

*2.2 Experiment participants and procedure*

The experiment spanned three months from September to November 2023, involving 77 participants aged 20 to 50 with driver's licenses. Of these, 39 were male (50.6%) and 38 female (49.4%), distributed across age groups as follows: 20s (27.9%), 30s (30.2%), 40s (23.3%), and 50s (18.6%).

The experimental procedure comprised three stages. Firstly, participants were completed a preliminary survey, followed by calibration for gaze measurement. Secondly, participants engaged in a 10-minute practice drive and a 6-minute simulation drive based on predefined scenarios. Finally, participants completed a post-experiment survey and received debriefing instructions.

*2.3 KDBQ Survey*

To analyze the driving behavior of participants before the experiment, we opted to employ the Korea Driver Behavior Questionnaire (KDBQ), a widely used scale that is a Korean translation of the Driver Behavior Questionnaire (DBQ). KDBQ is utilized to understand drivers' characteristics and behaviors, aiming to comprehend drivers in group units through their association with traffic accidents (Martinussen et al., 2013; Reason et al., 1990). The KDBQ consists of a total of 28 items, with each item rated on a 5-point scale (Lee & Kim, 2015).

In this experiment, the KDBQ was employed to assess drivers' characteristics and behaviors. Survey responses were statistically analyzed on a per-item basis, and factor analysis was conducted to extract principal components. This analysis aims to provide insights into the relationship between driver behavior and traffic accidents, facilitating a comprehensive understanding of drivers in the study cohort.

*2.4 Eye tracking data pre-processing and analysis*

The vehicle control data and eye-tracking data collected during the experiment must be accurately synchronized with timestamps. Each participant is assigned a unique ID, with vehicle control data capturing parameters such as vehicle speed and collision intensity, while eye-tracking data includes gaze coordinates, saccades, fixations, blinks, and pupil diameter.

Pre-processing of eye-tracking data involves calculating the frequency and duration of saccades, fixations, and blinks for each participant. Additionally, to facilitate comparison before and after accidents, the eye-tracking data for each driver must be segmented into 10-second intervals.

Following pre-processing, the eye-tracking data enables the characterization of eye movements for 10 seconds preceding and following accidents. These identified characteristics, analyzed alongside survey results, offer insights into the signals exhibited by drivers immediately before accidents occur.

## 3. Results

*3.1 Vehicle Accident Analysis*

*3.1.1 Accident Rate Analysis:*

Among the 77 drivers, 25 experienced accidents during the scenario drive, reflecting a 32.5% accident rate. Among these drivers, 18 were aged 40 or older, comprising 72% of the total, while 3 were in their 20s (12%) and 4 were in their 30s (16%).

Figure 3 illustrates the average values for the 10 seconds before and after the moment of collision for the 25 drivers involved in accidents. It demonstrates a tendency for collision intensity to peak simultaneously with the occurrence of accidents, followed by a rapid decrease within 10 seconds.

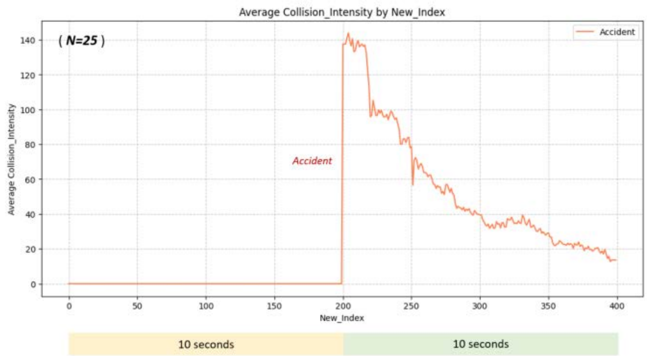



***Fig. 3.*** *Average collision intensity graph for drivers involved in accidents*

*3.1.2 Eye Tracking Data Analysis:*

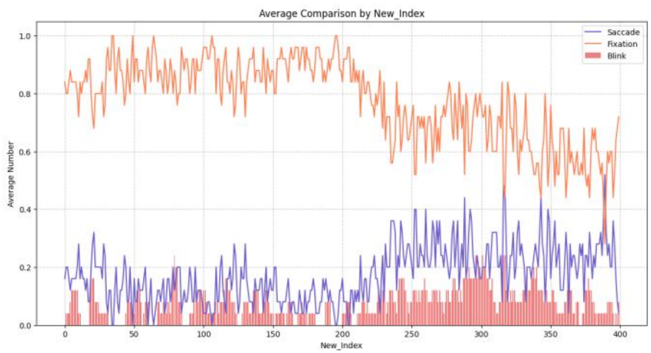



***Fig. 4.*** *Average comparative analysis of eye movement characteristics preceding and following accidents among drivers*

In Figure 4, a 20-second interval before and after collisions is depicted, with the collision moment marked at 200 on the x-axis. Examination of drivers' fixations reveals a sharp decrease in count after collisions, while saccades exhibit a pronounced increase. These patterns of eye movements serve as indicators reflecting drivers' psychological states following the occurrence of a collision event. Moreover, the occurrence of blinks before collisions appears sporadic and infrequent, whereas post-collision, the frequency of blinks increases, suggesting a heightened tendency for drivers to blink.

Additionally, Figure 5 provides insight into the pupil size of drivers involved in accidents. Pupil dilation begins approximately 10 seconds before collisions, implying the occurrence of emotionally charged events. Ultimately, the eye movement characteristics of drivers were analyzed concerning the point of collision in the misuse scenario.

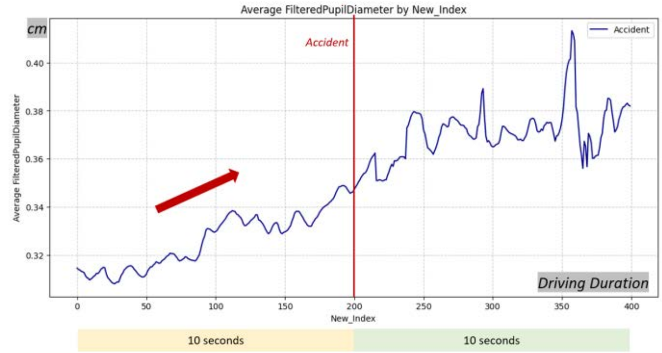



***Fig. 5.*** *Analysis of pupil diameter variations preceding and following collisions among involved drivers*

### *3.2 KDBQ Factor Analysis*

To explore the major factors underlying the 28 items of the KDBQ, exploratory factor analysis was conducted. Prior to factor analysis, items with extracted values below 0.5 from the communality analysis, such as KDBQ_Q3, were removed to ensure meaningful results. Consequently, factor analysis was carried out with the final set of 27 items, employing varimax rotation method for factor rotation. Principal component analysis was used for factor extraction, resulting in the identification of four factors considering the distribution of component matrices.

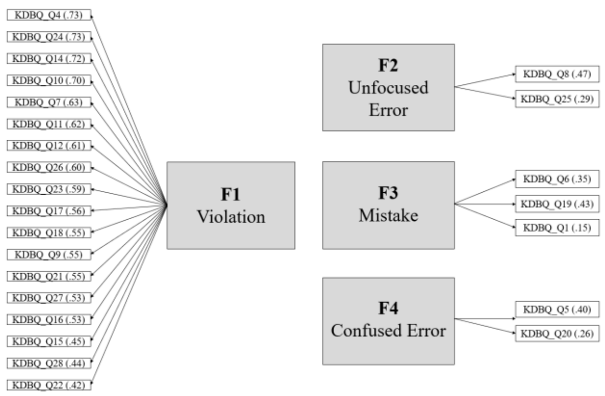



***Fig. 6.*** *Overall composition of extracted factors of KDBQ*

Reliability analysis of the factors extracted from the factor analysis demonstrates the validity of all 27 items, with Cronbach's alpha value reaching a very high level of 0.877. As shown in Figure 6, the first factor, with an Eigenvalue of 7.568, comprises 18 items primarily related to "Violation," indicating driving behaviors characterized by emotional driving and deliberate violations. The second factor, "Unfocused error," with an Eigenvalue of 2.074, encompasses behaviors such as failure to remember road conditions or check blind spots when changing lanes. The third factor, named "Mistake," includes errors such as failure to shift gears or misjudgement of vehicle speed. Lastly, the fourth factor, "Confused Error," involves behaviors like confusion about parking spaces or reversing in the wrong location.

These extracted factors serve as a foundation for classifying and specifying driver behaviors, which can be further linked to biosignals in secondary experiment analyses.

## 4. Discussion and conclusion

This study conducted vehicle simulator experiments based on scenarios where drivers misinterpret autonomous driving functionalities, aiming to extract and analyze driver behavior factors. A total of 77 participants took part in the misuse experiments, with 25 experiencing accidents during scenario execution. Factor analysis was conducted on the results of KDBQ, revealing four distinct factors. Additionally, analysis of eye-tracking data at the moment of accidents during the experiments allowed for the examination of pre- and post-accident eye movement characteristics. If such signals occur just before accidents in real-world scenarios, solutions can be provided based on the four factors identified through KDBQ analysis. Eye-tracking data can pinpoint driver reactions before and after accidents, with these reactions potentially linking to accident causation factors and subsequent response strategies. However, the study's generalizability is somewhat limited as it focused on a single scenario. In future work, we will expand the analysis to various misuse scenarios and provide a deeper understanding of driver monitoring in autonomous driving environments.


## 5. Acknowledgments

This work is supported by Korea Evaluation Institute of Industrial Technology(KEIT) grant funded by the Ministry of Trade, Industry and Energy(MOTIE, Korea) (No.20018248, Development of safety of the intended functionality from insufficiency of perception and decision making).


## References


Kim, J. et al. (2023). The Driver's Usability Analysis Study About Automated Driving Car. Korea Society for Emotion and Sensibility, *2023 Annual spring conference*, 121-121.

Martinussen, L. M., Hakamies-Blomqvist, L., Møller, M., Özkan, T. & Lajunen, T. Age, gender, mileage and the DBQ: The validity of the Driver Behavior Questionnaire in different driver groups. Accid. Anal. Prev. 52, 228–236 (2013).

Reason, J. T., Manstead, A., Stradling, S G., Baxter, J., Campbell, K. Errors and violations on the road – a real distinction. Ergonomics, 33 (10/11), 1315-1332 (1990).

Lee, C. H. & Kum, K. J. (2015). A Study on the Speeding Intention and Behaviors Based on a Driver Behavior Questionnaire. Journal of Korean Society of Transportation, 33(2), 159–169. https://doi.org/10.7470/jkst.2015.33.2.159

# Towards better subjective sleepiness ground truth data quality

Raimondas Zemblys [1*], Christer Ahlström[2,3], Anna Anund[2,4], Svitlana Finér [1]

[1] Smart Eye, AB, Gothenburg, Sweden
[2] Swedish National Road and Transport Research Institute, Linköping, Sweden
[3] Department of Biomedical Engineering, Linköping University, Linköping, Sweden
[4] Rehabilitation Medicine, Linköping University, Linköping, Sweden
[*]raimondas.zemblys@smarteye.se

**Abstract: The European Union's General Safety Regulations (GSR) mandates that all new vehicle types must be equipped with a driver sleepiness warning system from 2022. The test procedure used for evaluation and type approval of these systems relies heavily on subjective sleepiness ratings on the 9-point Karolinska Sleepiness Scale (KSS). This has raised concerns given that KSS is subjective and may be perceived and interpreted differently by different people. The objective of this paper is to investigate how different KSS training protocols affect the ratings, using data from eight different real-road studies. The results show that KSS ratings are more stable when participants are properly trained. The training could involve learning the scale by heart, for example by using the verbal anchors instead of the numbers during the experiment, and by extensively using the scale before the experiment starts, for example by filling in a sleep diary the days before the experiment. Everyone using KSS as a measure of sleepiness, regardless of if it is academia, companies, or regulatory bodies, must not underestimate the importance of calibrating the participants to enable accurate and absolute ratings.**

## 1. Introduction

The Karolinska Sleepiness Scale (KSS) is a subjective sleepiness rating scale (Table 1, Åkerstedt and Gillberg, 1990; Baulk et al., 2001) that is often used in driver sleepiness studies. KSS has been found to correlate with both objective and behavioural measures of sleepiness, and it is the measure of driver sleepiness that is least affected by inter-individual variations (Åkerstedt et al., 2014). However, the subjective feeling does not always reflect the actual sleepiness level and the anchored scale might be interpreted differently by different drivers. In studies investigating the development of sleepiness, these drawbacks may not be too important, but if KSS is used to develop and evaluate sleepiness detection systems, accurate and absolute ratings are essential. In the context of GSR and EuroNCAP, which both use KSS as the primary sleepiness metric in their stipulated requirements, all drivers must be trained to have the same understanding of the scale to be able to repeatedly quantify their state on multiple occasions and in different conditions and environments.

In this paper, eight datasets with sleepy drivers are analysed. The data were collected using different KSS training protocols to investigate how different protocols affect the ratings.

***Table 1** The Karolinska Sleepiness Scale*

| Description | Value |
|---|---|
| Extremely alert | 1 |
| Very alert | 2 |
| Alert | 3 |
| Rather alert | 4 |
| Neither alert nor sleepy | 5 |
| Some signs of sleepiness | 6 |
| Sleepy, no effort to keep awake | 7 |
| Sleepy, some effort to keep awake | 8 |
| Very sleepy, great effort to keep awake, fighting sleep | 9 |

## 2. Method

We divided the eight datasets into two categories (Table 2): "KSS-only instruction" and "Additional KSS instruction". In the first category (SleepEye, Mediator, NDDC), the original anchored KSS was used. In the latter (SmartEye-1, SmartEye-2, SmartEye-3, Fit2Drive, and "Expert-recordings"), the KSS labels were amended with additional explanations and examples on how to interpret and report KSS while driving. For example, KSS 6 could be described as "Some signs of sleepiness" and clarified by the description "This is the beginning of sleepiness, where you should be able to identify some sign of sleepiness. This can mean: Heavy eyes, some yawning, gaze being slower than normal (needing more time to process road signs, perhaps missing few of those), mental fatigue, etc. This does not have to be a progression from more alert states, for example, if sleep deprived or if last night's sleep was poor, one could be at this state from the very beginning of the drive". Further, a situation where KSS 6 typically occurs was provided as "Long drive, getting bored and slightly tired, experiencing some (very light) signs of sleepiness. Not yet a problem, can still drive and there is no need or urge to stop". The additional instructions were slightly different in the five "Additional KSS instruction" datasets, see Table 2.

The participants also followed slightly different training protocols on how to use the scale. In all eight studies, the scale was sent home to the participants a few days before the experiment, including any additional description of the labels. Upon arrival the test leader went through the scale one more time, explaining each level and label. In the SleepEye and Mediator datasets, the participants also filled in a sleep diary for three days before the study which forced the participants to practise using the scale. In the Fit2Drive dataset, the participants rated their sleepiness level by saying the full descriptive label rather than the number, with the intention to force the participants to reflect on the meaning of each level rather than just reporting the change from the previous rating.

***Table 2** Dataset information.*

| Dataset | Scenario | # drivers | Duration statistics* | KSS instructions |
|---|---|---|---|---|
| **"KSS-only instruction" datasets, multi-camera recordings** | | | | |
| SleepEye** (Fors et al., 2011) | Same highway route, Sweden | 17 | 17 recordings, 1.42-1.50 hours long (mean 1.44, 24 hours in total) | Verbal, visual (KSS scale on the steering wheel), sleep diary. |
| Mediator** (Ahlström et al., 2021) | Same highway route, Sweden | 66 | 67 recordings, 1.17-1.67 hours long (mean 1.45, 97 hours in total) | Verbal, sleep diary. |
| NDDC | Same highway route, short stop in between two drives, Sweden | 26 | 54 recordings, 0.73-1.59 hours long (mean 0.93, 50 hours in total) | Verbal. |
| **"Additional KSS instruction" datasets, single-camera recordings** | | | | |
| SmartEye-1 | Same highway route, US | 55 | 55 recordings, 2.23-3.76 hours long (mean 2.87, 158 hours in total) | Verbal amended with situational examples and a quiz, visual (coloured KSS scale). KSS 3 described as "9:30AM - the peak of your day and you feel as alert and awake as you ever do", KSS 6 as "2:30PM - You had a large lunch an hour ago, and are in the middle of a long meeting at work". |
| SmartEye-2 | Same highway route, US | 43 | 44 recordings, 1.34-3.40 hours long (mean 2.77, 122 hours in total) | Verbal with additional instructions, quiz. Additional instructions for each KSS level accompanied with an instruction "We do not want you to skip values". KSS 1–3 described as "very heightened and exaggerated state of being, something not often experienced"; KSS 4 as "Very attentive and alert driving", "Good cup of coffee after a full night of sleep", "This is the peak of your day and you feel as alert and awake as you ever do"; KSS 5 as "Causally attentive. One of the more common states for people driving", "Regular driving". |
| SmartEye-3 | Same highway route, US | 22 | 22 recordings, 1.00-1.09 hours long (mean 1.02, 22 hours in total) | Same as SmartEye-2. |
| Fit2Drive (Kircher et al., 2023) | ~20 min of city driving followed by one out of two highway routes, Sweden | 41 | 41 recordings, 2.15-3.00 hours long (mean 2.75, 113 hours in total) | Verbal with additional instructions, ratings with full label rather than numbers. Driving related examples for each KSS level. KSS 5 is described as "Regular relaxed driving, cruising for a long distance". |
| Expert-recordings | Naturalistic driving, Europe | 8 | 357 recordings, 0.15-5.11 hours long (mean 0.76, 272 hours in total) | Verbal with additional instructions. Instructions vary, might be self-taught or iteratively trained and updated. |
| Expert-recordings-subset | Naturalistic driving, Europe | 6 | 6 recordings, 1.07-5.11 hours long (mean 3.43, 21 hours in total) | Same as Expert-recordings |

* calculated using the difference between the first and the last KSS annotation
** only a subset of the original dataset with daytime manual drives on a highway were included

### *2.1 Datasets*

In all datasets KSS was retrospectively self-reported every 5th minute while driving. All datasets were recorded in a controlled field setting except the "Expert-recordings" dataset which was recorded in a naturalistic driving setting. The controlled experiments consisted of one trip per driver recorded during daytime on a predefined, mostly highway route with a test leader present in the vehicle. The "Expert-recordings" dataset included multiple drives per driver (up to 120 hours per driver) where each driver had extensive knowledge of KSS either from working on sleepiness detection algorithms or were extensively trained to perform self-evaluations. To enable comparisons with the other datasets we created an "Expert-recordings-subset" dataset where for each of the drivers we selected a single drive (the one closest to 3 hours).

Before the analysis the data were filtered. First, within each drive we removed data corresponding to single occurrences of KSS levels. Then we removed whole drives that had only one unique KSS level remaining, drives with a frame drop ≥5%, and drives where gaze data precision was ≥2.5 degrees in single-camera and ≥2.0 degrees in multi-camera recordings (precision measured as the median root mean square of the displacement between successive gaze position samples, S2S-RMS, over a 0.5-second sliding window; Hooge et al., 2018).

### *2.2 Metrics*

We looked at the data from two angles: whether the drivers were able to understand how to self-annotate KSS (operationalized by analysis of KSS distributions and rating dynamics), and whether they were able to evaluate themselves consistently multiple times (operationalized by stability evaluation of blink durations). Measure stability was estimated by first calculating the robust mean (mean of data with outliers removed using Tukey's fence method; Tukey, 1977) of blink durations within 5-minute sliding windows with an overlap of 15 seconds. Then all means were aggregated within each KSS level (the KSS value closest to the end of the window was chosen for each window) and the mean of interquartile ranges was calculated. The lower the stability metric, the more consistent the driver's blink durations are with respect to the subjective KSS ratings.

## 3. Results and discussion

### 3.1 Understanding of KSS

Reported KSS statistics in Figure 1 show that the KSS training protocol had a huge impact on the resulting KSS distribution. In general, most of the annotations were KSS 5 or 6 in the “Additional KSS instruction” datasets, while the KSS distributions were wider, including more KSS 3–4 and KSS 7–8 ratings in the “KSS-only” datasets. The SmartEye-1 dataset had the most distinctive distribution, which is most likely the consequence of KSS 3 being further described as “the peak of your day and you feel as alert and awake as you ever do”, thus biasing the ratings towards this level. Here drivers were also presented with a coloured KSS scale with very distinctive green, yellow, and red colouring of KSS levels 3, 7 and 9, which probably further reinforced the perception that “Alert” (KSS 3) was the “normal” level.

The Fit2Drive dataset, where drivers were asked to report definitions of KSS instead of numbers, had the largest amount of KSS 5 (“neither alert nor sleepy”) and, compared to similar datasets (SmartEye-2 and SmartEye-3), somewhat more KSS 4 (possibly indicating longer city driving) and fewer KSS 6–7 ratings (possibly due to a shorter amount of monotonous highway driving). Yet SmartEye-3 had many KSS values that indicated sleepiness (KSS 6–7) despite very short drives (~1 hour compared to ~3 hours in the other two datasets). A possible explanation is that different KSS training and rating protocols were used in the different datasets, where reporting the descriptive label leads to more accurate ratings than merely reporting the number. In “KSS-only instruction” datasets, the distributions of KSS were similar. In NDDC, ratings were shifted more to the right indicating more sleepy drivers, while SleepEye and Mediator datasets had more KSS 3–4 ratings. The reasons behind these differences are unclear and require further investigation.

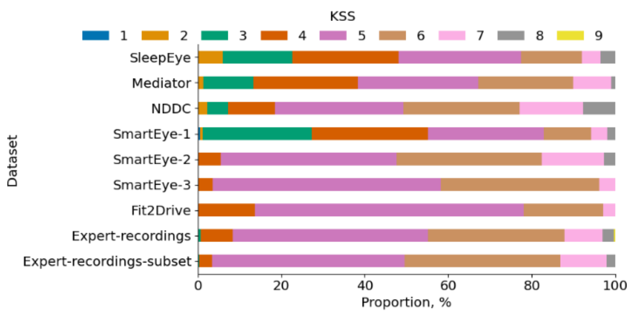



*Fig. 1. Distribution of KSS ratings in all datasets.*

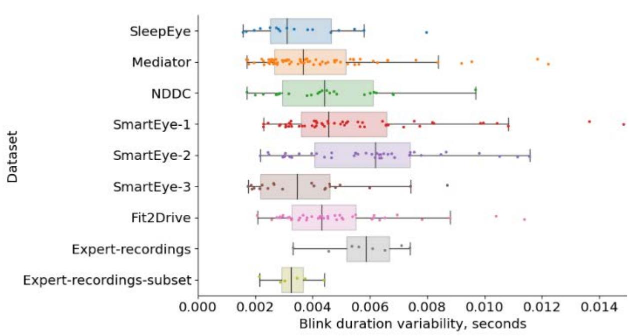



*Fig. 2. Blink duration stability in all datasets (lower is better; trimmed at 0.015 seconds for illustration purposes). Dots represent stability for each driver.*

Looking into KSS dynamics, the “Additional KSS instruction” datasets had fewer KSS level changes compared to the “KSS-only instruction” datasets. Most of the changes were ±1 KSS level in all datasets. Most of the -2 KSS level changes happened in the middle or the end of the drive likely indicating (i) a turn-around point, (ii) coming back to city-like driving, or (iii) end of the drive. Most of the +2 and +3 KSS levels changes happened in the beginning and up to the middle of the drives, indicating that drivers became sleepier early on in their drives.

### 3.2 Measure stability

Figure 2 illustrates that blink duration was most stable in the SleepEye, Mediator, SmartEye-3, and “Expert-recordings-subset” datasets. For the SleepEye, Mediator and “Expert-recordings-subset” datasets, improved stability is probably due to more rigorous training of the participants in how to use the scale, whereas for SmartEye-3, high stability is likely due to short drives with less chances for the blink duration to vary within the same KSS level. The differences between SmartEye-2 and SmartEye-3, and also between “Expert-recordings-subset” and “Expert-recordings”, points in the same direction, i.e. that drive duration and varying driving environments results in worsened measure stability. Overall, the results show that blink duration stability improves when (i) a sleep diary is used, and (ii), when description-based KSS annotations are made. Both approaches force the participants to practise and learn the scale by heart.

### 3.3 General discussion

Our analyses shows that KSS training instructions might have an impact on the reported KSS levels. Similar training instructions resulted in similar KSS distributions across different datasets and less KSS level changes throughout the drive(s). However, additional instructions might also affect the distribution (see SmartEye-1 in Figure 1). Similar observations were made in the NDDC dataset where no additional instructions explaining KSS were used, and moreover, drivers had very limited exposure to the scale before the drive (as opposed to e.g. using a sleep diary before the drive as in the SleepEye and Mediator datasets). This indicates that drivers who have not learned the scale well enough instead develops an own understanding of the scale, which leads to confusion and less accurate ratings.

Overall, we conclude that description-based ratings (as used in the Fit2Drive dataset) reflects driving conditions best and align state understanding across the drivers better. Such training resulted in reasonable KSS distributions and dynamics, while also giving rather stable blink durations over different KSS levels. We believe that using numbers to report KSS levels might hide the non-linear nature of the KSS (i.e., that physiological and behavioural sleepiness indicators increase exponentially with a linear increase in KSS, see Åkerstedt et al., 2014). However more controlled studies are needed to remove the many confounds we had in and between our datasets (different routes, driving conditions, driver backgrounds, etc.). We believe that rigorous learning of the scale, in combination with a standardised data collection protocol, is needed to get replicable results. All involved parties, regardless if it is academia, companies, or regulatory bodies, should follow the same procedure in collecting subjective KSS data.

**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.

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. International journal of neuroscience, 52(1-2), 29-37.

Åkerstedt, T., Anund, A., Axelsson, J., & Kecklund, G. (2014). Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. Journal of sleep research, 23(3), 242-254.

Baulk, S. D., Reyner, L. A., & Horne, J. A. (2001). Driver sleepiness—evaluation of reaction time measurement as a secondary task. Sleep, 24(6), 695-698.

Fors, C., Ahlström, C., Sörner, P., Kovaceva, J., Hasselberg, E., Krantz, M., ... & Anund, A. (2011). Camera-based sleepiness detection: final report of the project SleepEYE. Statens väg-och transportforskningsinstitut.

Hooge, I. T., Niehorster, D. C., Nyström, M., Andersson, R., & Hessels, R. S. (2018). Is human classification by experienced untrained observers a gold standard in fixation detection?. Behavior Research Methods, 50, 1864-1881.

Kircher, K., Finér, S., Zemblys, R., & Ahlström, C. (2023). Fit2Drive.

Tukey, J. W. (1977). Exploratory data analysis (Vol. 2, pp. 131-160).

# The influence of cognitive load on driver performance in automated driving contexts: a scoping review in progress

Hao Qin [1*], Natasha Merat [1], Rafael Cirino Gonçalves [1]

[1] Institute for Transport Studies, University of Leeds, Leeds, United Kingdom
[*]tshq@leeds.ac.uk

**Abstract: Increasing vehicle automation (SAE Level 2 & 3) makes drivers prone to engagement with non-driving-related activities. Activities such as hands-free phone conversations may expose drivers to cognitive load and interfere with their safety-critical performance after a takeover. While the effect of cognitive load in manual driving conditions has been well-documented, its influence in automated driving is less well understood. Given the growing number of studies in this ear, we conducted a scoping review focusing on the influence of cognitive load on driver performance in automated driving. As the examination of multiple databases is still in progress, this paper presents our preliminary results, based on Scopus database search. Nine papers met the requirement and were included in the review. Results showed that cognitive load affects multiple aspects of driver performance, including monitoring-related performance, such as eye-movements, peripheral vision, situation awareness, and auditory perception. Quality of a takeover, but not its timing, is also found to be affected cognitive load. We identified several limitations of the reviewed literature, including the small volume of papers, biased experimental scenarios, and a lack of real driving studies. More research is needed to address these limitations, to gain a better understanding of the effect of cognitive load on performance in automated driving.**

## 1. Introduction

Driver distractions can be classified into three main types, including visual, manual, and cognitive distractions (Vegega et al., 2013). Cognitive distraction commonly refers to a "withdrawal of attention from the driving task" due to competing activities (Engstrom et al., 2017; Young et al., 2003). The amount of cognitive resource demanded by these activities is normally termed cognitive load (Engstrom et al., 2017). In this paper, we specifically focus on the cognitive load induced by cognitive distractions without visual and manual components, such as the demand required by hands-free phone conversations.

The influence of cognitive load on manual driving performance is well-documented. For instance, cognitive load causes gaze concentration towards the road center and reduces attention to other safety-related areas, such as wing mirrors and off-path hazards (Broadbent et al., 2023; Nilsson et al., 2020). Drivers' control of the vehicle, such as maintaining longitudinal speed, is also found to be affected by cognitive load (Ma et al., 2023). The "cognitive control hypothesis" (Engstrom et al., 2017) suggests that cognitive load impairs driving tasks that rely on cognitive control, leaving automatized responses unaffected (e.g., braking response to looming).

The impact of cognitive load in automated driving remains less well understood. Cognitive distractions, such as hands-free phone conversations, are prevalent in automated driving (Carsten et al., 2012; Reagan et al., 2021). Studies are demonstrating that cognitive load compromises performance in automated driving, such as reducing drivers' ability to detect peripheral hazards or safely take over control of the driving task, when the system reaches its limits (Choi et al., 2020; Merat et al., 2014; Yang et al., 2022). However, as far as we are aware, there is currently a shortage of literature that reviews and synthesizes the effect of cognitive load on drivers during different levels of automated driving.

To fill this gap, we conducted a scoping review, specifically focusing on the influence of cognitive load on drivers' attention and driving performance in automated driving. We focused on studies involving SAE Level 2 and 3 automation, where the vehicle is responsible for the driving task if automation is activated, but human intervention may be needed when the system reaches a limitation (SAE International, 2021).

## 2. Method

The review process followed the guidance published by Peters et al. (2015). Two clusters of search terms were used for the literature search: (i) "driving" or "driver" and "automated", "autonomous", "automation", or "autopilot"; (ii) "cognitive" or "mental", and "load", "workload" "distraction", or "demand". The Title, Abstract, and Keywords of outputs within four databases, including IEEE Xplore, ACM Digital Library, SAE Mobilus and Scopus, were searched. We only considered peer-reviewed journal articles and conference papers written in English and did not include a year limit.

Based on the scoping review, articles were excluded if they did not (i) distinguish between cognitive distraction and other types of distraction; (ii) compare different levels of cognitive load; (iii) focus on drivers' attention or driving performance; (iv) examine SAE Level 2 or 3 automation. Articles were first screened by Title and Abstract. Those that passed these criteria were further examined in full text. As the examination of additional databases is currently in progress, we only present results identified by the Scopus database, reporting on results from a total of 551 articles. Only nine articles passed the above screening process and are discussed further below.

## 3. Results & Conclusions

### *3.1 Influence of cognitive load on driver monitoring-related activities*

Five out of the nine papers examined the impact of cognitive load on drivers' monitoring-related activities, including eye-movements, peripheral vision, situation awareness and auditory perception. Studies on eye-movement yielded inconsistent results (N=2). Heikoop et al. (2018) found that engagement in a verbal 2-back task did not influence drivers' horizontal gaze dispersion during automated driving, indicating no gaze concentration. On the other hand, Wu et al. (2019) investigated saccadic behaviours and found an increased percentage of small saccades under cognitive load during automated driving.

Two studies found an effect of cognitive load on drivers' detection of peripheral targets. Van Winsum (2019) found that engagement in a backwards counting task delayed drivers' reaction time in a peripheral detection task during automated driving. Similarly, Yang et al. (2022) found lower detection rate of peripheral targets under cognitive load during real-world automated driving.

Drivers' situation awareness during automated driving also appears to be affected by cognitive load (N=1). Heikoop et al. (2018) used a think-aloud protocol (participants were required to verbalize any thoughts during the drive) to probe drivers' situation awareness during an autonomous car platoon study and found diminished awareness of driving-related situations among cognitively loaded drivers.

Using event-related potentials, Van der Heiden et al. (2022), demonstrated that cognitive load, induced by a non-visual cognitive task, reduced brain sensitivity to external auditory stimuli during automated driving.

### *3.2 Influence of cognitive load on driver takeover performance*

The effect of cognitive load on the timing and quality of takeovers has also been studied (N=5). Findings suggest that cognitive load does not interfere with takeover time. Louw et al. (2017) found that cognitive load induced by a verbal 1-back task did not affect the time drivers take over vehicle control to avoid collision with a braking lead vehicle. The onset of obstacle avoidance behaviour, such as braking or steering reaction time, are also found to be immune to cognitive load (Choi et al., 2020; Louw et al., 2017; Lu et al., 2021; Wu et al., 2019).

On the other hand, drivers' takeover quality is potentially affected by cognitive load (N=2). Choi et al. (2020) reported that cognitive load delayed the completion of a lane crossing when attempting to avoid a lead obstacle after takeover. However, when it comes to crash probability, no effect of cognitive load is found, as most drivers were able to avoid collisions, regardless of cognitive load (Choi et al., 2020; Louw et al., 2017). However, results are somewhat contradictory. For example, Lu et al. (2021) found that cognitive load improved takeover quality, demonstrated by increased minimum time-to-collision during lead vehicle avoidance. This difference in findings may be due to the different automation levels used by the two studies. While Choi et al. (2020) examined Level 2 automation, where drivers were required to monitor the road and an additional cognitive task may have caused overload, Lu et al. (2021) investigated Level 3 automation, where monitoring was not required. It can be argued that the additional load from the cognitive task prevented drivers from underload and improved their takeover performance, although further work needs to be done in this context.

Based on the above findings, there is some evidence that cognitively loading tasks may influence driver performance during automated driving. However, considering the limited volume of studies, this knowledge remains to be validated by more work. It is also worth noting that all takeover scenarios used in these studies were limited to unexpected lead obstacle avoidance. It has been argued that drivers respond automatically to such unpredictable obstacles due to looming, which requires little cognitive control and is immune to cognitive load (Engstrom et al., 2017; Louw, Markkula, et al., 2017). This possibly explains the absence of an effect from cognitive load on takeover timing. Further work in this area, examining the effect of cognitive load on other types of takeover scenario is therefore warranted. A limitation of the papers outlined above is also that all but one was based on driving simulator studies, which may results in different types and speed of response, compared to real-world settings (Gemonet et al., 2021). It is hoped that a more comprehensive overview of additional studies will be reported in time for the conference in October, when more databases are used for the scoping review.

## 4. Acknowledgments

The first author is funded by a China Scholarship Council-University of Leeds Scholarship.

**References**

Broadbent, D. P., D'Innocenzo, G., Ellmers, T. J., Parsler, J., Szameitat, A. J., & Bishop, D. T. (2023). Cognitive load, working memory capacity and driving performance: A preliminary fNIRS and eye tracking study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *92*, 121–132. https://doi.org/10.1016/j.trf.2022.11.013

Carsten, O., Lai, F. C. H., Barnard, Y., Jamson, A. H., & Merat, N. (2012). Control Task Substitution in Semiautomated Driving: Does It Matter What Aspects Are Automated? *Human Factors*, *54*(5), 747–761. https://doi.org/10.1177/0018720812460246

Choi, D., Sato, T., Ando, T., Abe, T., Akamatsu, M., & Kitazaki, S. (2020). Effects of cognitive and visual loads on driving performance after take-over request (TOR) in automated driving. *Applied Ergonomics*, *85*, 103074. https://doi.org/10.1016/j.apergo.2020.103074

Engstrom, J., Markkula, G., Victor, T., & Merat, N. (2017). Effects of cognitive load on driving performance: The cognitive control hypothesis. *Human Factors*, *59*(5), Article 5.

Gemonet, E., Bougard, C., Masfrand, S., Honnet, V., & Mestre, D. R. (2021). Car drivers coping with hazardous events in real versus simulated situations: Declarative, behavioral and physiological data used to assess drivers' feeling of presence. *PLOS ONE*, *16*(2), e0247373. https://doi.org/10.1371/journal.pone.0247373

Heikoop, D. D., de Winter, J. C. F., van Arem, B., & Stanton, N. A. (2018). Effects of mental demands on situation awareness during platooning: A driving simulator study.

*Transportation Research Part F: Traffic Psychology and Behaviour*, *58*, 193–209. https://doi.org/10.1016/j.trf.2018.04.015

Louw, T., Madigan, R., Carsten, O., & Merat, N. (2017). Were they in the loop during automated driving? Links between visual attention and crash potential. *Injury Prevention*, *23*(4), 281–286.

Louw, T., Markkula, G., Boer, E., Madigan, R., Carsten, O., & Merat, N. (2017). Coming back into the loop: Drivers' perceptual-motor performance in critical events after automated driving. *Accident Analysis & Prevention*, *108*, 9–18. https://doi.org/10.1016/j.aap.2017.08.011

Lu, G., Zhai, J., Li, P., Chen, F., & Liang, L. (2021). Measuring drivers' takeover performance in varying levels of automation: Considering the influence of cognitive secondary task. *Transportation Research Part F: Traffic Psychology and Behaviour*, *82*, 96–110. https://doi.org/10.1016/j.trf.2021.08.005

Ma, J., Wu, Y., Rong, J., & Zhao, X. (2023). A systematic review on the influence factors, measurement, and effect of driver workload. *Accident Analysis & Prevention*, *192*, 107289. https://doi.org/10.1016/j.aap.2023.107289

Merat, N., Jamson, A. H., Lai, F. C. H., Daly, M., & Carsten, O. M. J. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 274–282. https://doi.org/10.1016/j.trf.2014.09.005

Nilsson, E. J., Victor, T., Ljung Aust, M., Svanberg, B., Lindén, P., & Gustavsson, P. (2020). On-to-off-path gaze shift cancellations lead to gaze concentration in cognitively loaded car drivers: A simulator study exploring gaze patterns in relation to a cognitive task and the traffic environment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *75*, 1–15. https://doi.org/10.1016/j.trf.2020.09.013

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, *13*(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050

Reagan, I. J., Teoh, E. R., Cicchino, J. B., Gershon, P., Reimer, B., Mehler, B., & Seppelt, B. (2021). Disengagement from driving when using automation during a 4-week field trial. *Transportation Research Part F: Traffic Psychology and Behaviour*, *82*, 400–411. https://doi.org/10.1016/j.trf.2021.09.010

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/

Van der Heiden, R. M. A., Kenemans, J. L., Donker, S. F., & Janssen, C. P. (2022). The Effect of Cognitive Load on Auditory Susceptibility During Automated Driving. *Human Factors*, *64*(7), 1195–1209. https://doi.org/10.1177/0018720821998850

Van Winsum, W. (2019). Optic Flow and Tunnel Vision in the Detection Response Task. *Human Factors*, *61*(6), 992–1003. https://doi.org/10.1177/0018720818825008

Vegega, M., Jones, B., & Monk, C. (2013). *Understanding the effects of distracted driving and developing strategies to reduce resulting deaths and injuries: A report to congress* (Report No. DOT HS 812 053).

Wu, Y., Kihara, K., Takeda, Y., Sato, T., Akamatsu, M., & Kitazaki, S. (2019). Assessing the Mental States of Fallback-Ready Drivers in Automated Driving by Electrooculography. *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 4018–4023. https://doi.org/10.1109/ITSC.2019.8917029

Yang, S., Wilson, K., Roady, T., Kuo, J., & Lenné, M. G. (2022). Beyond gaze fixation: Modeling peripheral vision in relation to speed, Tesla Autopilot, cognitive load, and age in highway driving. *Accident Analysis & Prevention*, *171*, 106670. https://doi.org/10.1016/j.aap.2022.106670

Young, K., Regan, M., & Hammer, M. (2003). *Driver Distraction: A Review of the Literature*. Rep. No. 206. Victoria, Australia: Monash University Accident Research Centre.

# Decoding of Event-Related Potentials Through Steering Wheel Movements

Shahar Aberbach-Goodman[1], Eldad Izhak Hochman [1,2]

[1] CorrActions Ltd., Concino 16, Tel-Aviv, Israel
[2] Department of Psychology and Education, The Open University of Israel, Derekh ha-Universita 1, Ra'anana, Israel

[*]eldad@corractions.com

**Abstract: EEG is considered the most reliable measure for identifying a driver's cognitive state. Here, we present a new algorithm to identify the driver's state, using steering wheel movements underlying evoked potentials. First, we coupled error-related evoked potentials with unique corrective sub-movements recorded from a steering wheel in a driving simulator. Next, drivers' blood alcohol counts (BAC), and fatigue (KSS) were parametrically manipulated during real driving. Finally, a classification model was trained based on the corrective sub-movements to distinguish between different driving impairments (three BAC and two fatigue levels). Our results revealed sub-movement features uniquely sensitive to either BAC or fatigue, in addition to features sensitive to the accumulative effect of both impairments. Based on the corrective sub-movements, we are currently further developing a new model for detecting cognitive distraction while driving.**

## 1. Introduction

**It is commonly assumed that brain electrical activity recorded using an electroencephalogram (EEG) is the most effective way to detect the cognitive state of a car driver in real time (Fouad, 2023). However, this method is less applicable for driving scenarios outside the laboratory or clinic.**

**Performance monitoring is the process by which the brain forms predictions and monitors action outcomes (Ullsperger et al., 2014). Action outcomes are processed regardless of whether the information originates from internal (e.g., proprioceptive) or external (e.g., visual) sensory channels (Müller et al., 2005). Critically, previous work shows that the error-related EEG components of performance monitoring appear continuously while driving (Zhang et al., 2015).**

**Several theoretical models indicate that error-related EEG components reflect processes of motor compensation for making an error (Hochman et al., 2015; Hochman, Orr, et al., 2014; Hochman, Vaidya, et al., 2014). Accordingly, a series of studies show that these EEG components are associated with unique activation patterns in the cerebral motor cortex responsible for muscle activation and the production of movement (Even-Chen et al., 2018; Hochman et al., 2009).**

**In the current study, subjects drove a simulator while both EEG and steering wheel movements were measured. Using machine learning (ML) we built an index of unique corrective sub-movements that only appear when an error is detected in the brain. The index allowed us to identify when error detection processes occur, without the need for EEG recordings. In several follow-up experiments, we collected steering wheel angle data during real driving under intoxication and fatigue conditions. Based on the corrective sub-movements, our ML models distinguished between three levels of intoxication and two levels of fatigue. The results show significant detection of the different levels of impairments with no miss classification of one impairment as another. Furthermore, we were able to identify the cumulative effect of intoxication and fatigue on cognition.**

## 2. Method

### *2.1 EEG*

*Participants:* 52 participants (18 women). Age, 23-46 years. *Protocol:* EEG was recorded using a 64-channel cap with 10-20 design, digitized at 2048 Hz. The steering wheel angle was sampled at 200 Hz with a resolution of 0.1 degrees. Participants drove a 15-minute driving simulation of urban and highway scenarios.

### *2.2 Intoxication*

*Participants:* 127 participants (59 women). Age, 24-64 years. *Protocol:* Across all studies, participants drove vehicles from seven manufacturers on six closed tracks. The steering wheel angle was sampled at 200 Hz with a resolution of 0.1 degrees, obtained from the vehicle CAN bus. BAC was manipulated within subjects by drinking 40% alcohol 15 minutes before driving. Alcohol levels of 0, 0.05, and 0.08 BAC were verified by Breathalyzer. The drinking protocol followed Watson et al., 1981

### 2.3 Fatigue

*Participants:* 30 participants (12 women). Age, 26-46 years. *Protocol:* Closed-track driving. Steering wheel angle sampled at 200 Hz with a resolution of 0. 1 degree, obtained from the vehicle CAN bus. 28 hours of sleep deprivation. 3 hours driving. Karolinska sleepiness scale (KSS) reported every 15 minutes.

## 3. Results

### 3.1 Coupling corrective sub-movement with error-related negativity:

Sub-movements extracted from steering acceleration profiles, defined as samples exceeding both their neighboring values and zero. EEG recordings from FCz electrode were divided into one-second epochs, each centered on a single acceleration peak. A Sub-movement was labeled as corrective if its EEG epoch contained an error-related negativity (ERN). An ML model was trained to distinguish corrective from non-error-related sub-movements. Half the data was used for training and the other half for testing, showing 98% accuracy.

### 3.2 Intoxication and fatigue

Steering wheel movements (steering angle) were used to distinguish three BAC levels and separately, two levels of fatigue (below and above KSS = 7). For BAC, our model reached FP = 0%, and TP = 100%. Fatigue detection showed FP = 3% and TP = 83%. The results for both impairments showed generalization across different participants.

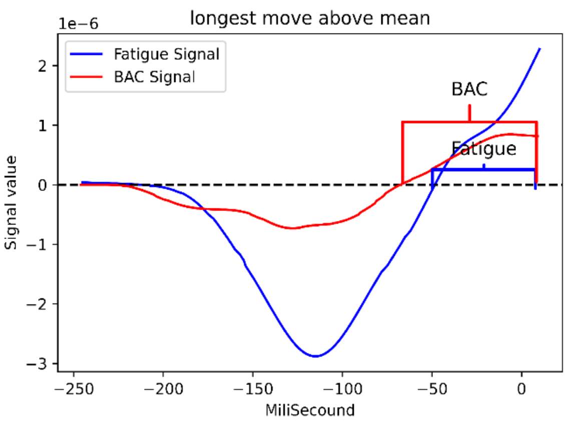



***Fig. 1.*** *Feature sensitive to BAC - The length of the longest sub-signal higher than the averaged signal.*
*A longer signal is obtained for BAC compared to fatigue. BAC > 0.08%, Fatigue KSS > 7.*

### 3.3 Distinguishing intoxication from fatigue:

Based on the steering wheel angle, the model was required to distinguish between fatigue and intoxication. The level of accuracy was 100%. Three types of features have been identified: intoxication sensitive, fatigue sensitive, and sensitive to both.

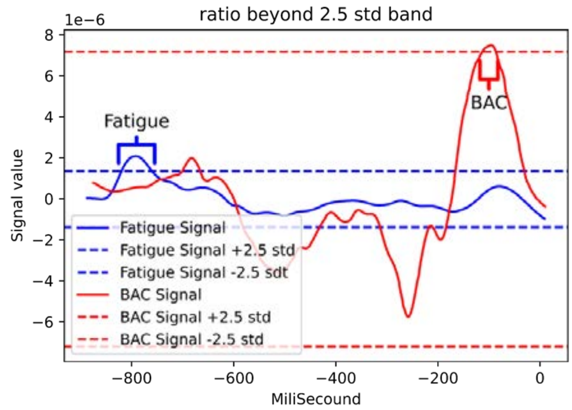



***Fig. 2.*** *Feature sensitive to fatigue- The ratio of signal higher than 2.5 STD.*
*Fatigue (KSS > 7) condition has a higher ratio compared to BAC > 0.08%).*

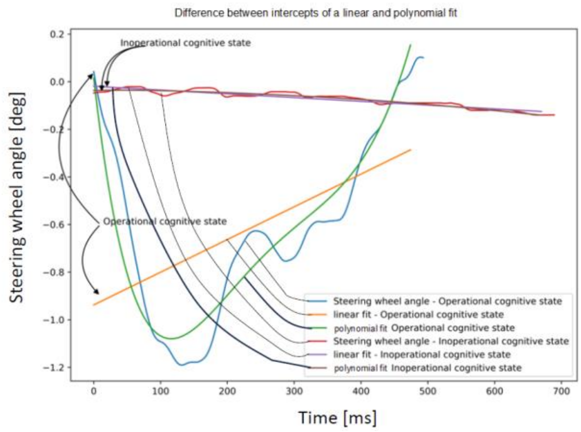



***Fig. 3.*** *Example of a feature sensitive to both BAC and fatigue.*
*For a driver presenting KSS = 7 and BAC = 0.05, the difference between intercepts of a linear and polynomial fit is similar to a driver presenting BAC above the legal threshold (0.08).*

***Table 1.*** *Identification of BAC at three thresholds*

| Track | BAC | Accuracy | True Positives | False Positives |
|---|---|---|---|---|
| 1 | 0.1 | 91% | 81% | 0% |
| 2 | 0.1 | 87% | 80% | 0% |
| 3 | 0.08 | 100% | 100% | 0% |
| 4 | 0.05 | 100% | 100% | 0% |
| 5 | 0.05 | 95% | 90% | 0% |
| 6 | 0.05 | 95% | 75% | 0% |

## 4. Discussion

By coupling EEG recordings and steering wheel movements in a driving simulator, we identified sub-movements that only appear when performance monitoring related evoked potentials emerge in the brain. Then, through parametric manipulation of intoxication and fatigue in real driving and using machine learning, we identified the crossing of 3 BAC thresholds: 0.05, 0.08, and 0.1 at 87-100% accuracy levels and the crossing of KSS=7 level, at 90% accuracy, all within 7 minutes of recording. Our model's performance is superior to previously reported analyses of real-road driving steering wheel sub-movements (Sikander & Anwar, 2019). Critically, our results demonstrate the robustness of the method for data obtained from a variety of vehicle models and natural driving routes.

Moreover, we extracted sub-movements features that respond uniquely to intoxication, fatigue or to both. This capability is unique to sub-movements analysis. Since error-related potentials are similarly affected by fatigue and intoxication (Ridderinkhof et al., 2002; Scheffers et al., 1999), the impairments are less distinguishable using EEG .

It is possible that intoxication and fatigue differentially effect sub-movements patterns since each of these factors uniquely effect the planning and execution processes of movements and consequently error-related compensation processes (Noël et al., 2010; Rozand et al., 2015).

This capability has implications for driver monitoring. First, the regulation requires identification of a specific problem source (National Highway Traffic Safety Administration, 2023). Second, when a specific source is identified, a specific solution can be proposed whose effectiveness is valid only or mainly for that specific factor (May & Baldwin, 2009). Third, we assume that the features sensitive to both fatigue and intoxication will be sensitive to almost any other factor affecting driver cognition. Thus, a model can be built for detection of a general deterioration in driver's operative ability. Here, a key challenge is determining the warning threshold. A model that relies on features sensitive to both alcohol and fatigue provides an elegant solution. In a situation where the driver is slightly drunk and slightly tired, the unique indicators of intoxication or fatigue do not detect threshold crossing. However, the cumulative effect of intoxication and fatigue may deteriorate the driver's operational state to a level that requires an alert. Here, the features sensitive to both alcohol and fatigue will be more strongly affected than the exclusive features. Thus, unlike the exclusive ones, they will reach the regulatory alert threshold level for alcohol or fatigue and an intervention will be issued.

Existing solutions for detecting distraction while driving have great difficulty dealing with situations such as cognitive overload, in which the driver's mind is attentive to road conditions but is unable to process them effectively (Kashevnik et al., 2021). The cerebral performance monitoring system is sensitive to cognitive load (Krigolson et al., 2012) and therefore it is likely that the solution presented in this article may be generalized to situations of cognitive load. Using the methods presented here, we are currently developing a model that will detect cognitive load while driving.

Our methodology is not limited to steering angle analysis. We have conducted a variety of experiments in which driver and passenger movements were monitored using seat sensors or in-cabin radar. Moreover, because motion sensors exist in almost any digital device, our methodology can be implemented in various fields such as gaming, defense, and health.

## 5. Conclusions

The present study shows that ERPs can be monitored by analyzing body movements. This technique can be used to identify a driver's cognitive operational state for activation of vehicle safety intervention in real time.

## References

Even-Chen, N., Stavisky, S. D., Pandarinath, C., Nuyujukian, P., Blabe, C. H., Hochberg, L. R., Henderson, J. M., & Shenoy, K. V. (2018). Feasibility of automatic error detect-and-undo system in human intracortical brain-computer interfaces. *IEEE Transactions on Biomedical Engineering*, *65*(8), 1771–1784. https://doi.org/10.1109/TBME.2017.2776204

Fouad, I. A. (2023). A robust and efficient EEG-based drowsiness detection system using different machine learning algorithms. *Ain Shams Engineering Journal*, *14*(3), 101895. https://doi.org/10.1016/j.asej.2022.101895

Hochman, E. Y., Eviatar, Z., Breznitz, Z., Nevat, M., & Shaul, S. (2009). Source localization of error negativity: Additional source for corrected errors.

*NeuroReport*, *20*(13), 1144–1148. https://doi.org/10.1097/WNR.0b013e32832f84ed

Hochman, E. Y., Orr, J. M., & Gehring, W. J. (2014). Toward a More Sophisticated Response Representation in Theories of Medial Frontal Performance Monitoring: The Effects of Motor Similarity and Motor Asymmetries. *Cerebral Cortex*, *24*(2), 414–425. https://doi.org/10.1093/cercor/bhs323

Hochman, E. Y., Vaidya, A. R., & Fellows, L. K. (2014). Evidence for a Role for the Dorsal Anterior Cingulate Cortex in Disengaging from an Incorrect Action. *PLoS ONE*, *9*(6), e101126. https://doi.org/10.1371/journal.pone.0101126

Hochman, E. Y., Wang, S., Milner, T. E., & Fellows, L. K. (2015). Double dissociation of error inhibition and correction deficits after basal ganglia or dorsomedial frontal damage in humans. *Neuropsychologia*, *69*, 130–139. https://doi.org/10.1016/j.neuropsychologia.2015.01.023

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

Krigolson, O. E., Heinekey, H., Kent, C. M., & Handy, T. C. (2012). Cognitive load impacts error evaluation within medial-frontal cortex. *Brain Research*, *1430*, 62–67. https://doi.org/10.1016/j.brainres.2011.10.028

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

Müller, S. V., Möller, J., Rodriguez-Fornells, A., & Münte, T. F. (2005). Brain potentials related to self-generated and external information used for performance monitoring. *Clinical Neurophysiology*, *116*(1), 63–74. https://doi.org/10.1016/j.clinph.2004.07.009

National Highway Traffic Safety Administration. (2023). Advanced Impaired Driving Prevention Technology (ANPRM 2127-AM50). U.S. Department of Transportation. https://www.regulations.gov/document/NHTSA-2022-0079-0015

Noël, X., Tomberg, C., Verbanck, P., & Campanella, S. (2010). The Influence of Alcohol Ingestion on Cognitive Response Inhibition and Error Processing. *Journal of Psychophysiology*, *24*(4), 253–258. https://doi.org/10.1027/0269-8803/a000039

Ridderinkhof, K. R., De Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., & Band, G. P. H. (2002). Alcohol Consumption Impairs Detection of Performance Errors in Mediofrontal Cortex. *Science*, *298*(5601), 2209–2211. https://doi.org/10.1126/science.1076929

Rozand, V., Lebon, F., Papaxanthis, C., & Lepers, R. (2015). Effect of mental fatigue on speed–accuracy trade-off. *Neuroscience*, *297*, 219–230. https://doi.org/10.1016/j.neuroscience.2015.03.066

Scheffers, M. K., Humphrey, D. G., Stanny, R. R., Kramer, A. F., & Coles, M. G. H. (1999). Error-related processing during a period of extended wakefulness. *Psychophysiology*, *36*(2), 149–157. https://doi.org/10.1111/1469-8986.3620149

Sikander, G., & Anwar, S. (2019). Driver Fatigue Detection Systems: A Review. *IEEE Transactions on Intelligent Transportation Systems*, *20*(6), 2339–2352. https://doi.org/10.1109/TITS.2018.2868499

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of Performance Monitoring and Adaptive Behavior. *Physiological Reviews*, *94*(1), 35–79. https://doi.org/10.1152/physrev.00041.2012

Watson, P. E., Watson, I. D., & Batt, R. D. (1981). Prediction of blood alcohol concentrations in human subjects. Updating the Widmark Equation. *Journal of Studies on Alcohol*, *42*(7), 547–556. https://doi.org/10.15288/jsa.1981.42.547

Zhang, H., Chavarriaga, R., Khaliliardali, Z., Gheorghe, L., Iturrate, I., & Millán, J. D. R. (2015). EEG-based decoding of error-related brain activity in a real-world driving task. *Journal of Neural Engineering*, *12*(6), 066028. https://doi.org/10.1088/1741-2560/12/6/066028

# A Novel Paradigm for Identifying Eye-Tracking Metrics Associated with Cognitive Control During Driving Through MEG Neuroimaging

Thomas Seacrist[1,2*], Elizabeth E. Walshe[1], Shukai Cheng[1], Emily Brown[1], Charlotte Birnbaum[4,5], Victoria Kaufman[4,5], Flaura K. Winston[1,3], William Gaetz[1,4,5]

[1]Center for Injury Research and Prevention, Children's Hospital of Philadelphia, Philadelphia, PA, USA
[2]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA
[3]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[4] Lurie Family Foundations' MEG Imaging Center, Department of Radiology, Children's Hospital of Philadelphia, Philadelphia, PA, USA
[5] Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
[*]seacrist@chop.edu

**Abstract**: Recent developments in eye-tracking and neuroimaging are paving the way for novel approaches to understanding cognitive errors during driving. If we can link fundamental neurocognitive capacities that are important for situational awareness with driver behaviors that are detectable via in-vehicle eye-tracking, we may be able to detect deficits in drivers' cognitive function in real-time. To this end, this study tests the utility of adding eye tracking recording to a recently validated novel paradigm that combines simulated driving with simultaneous magnetoencephalography (MEG) neuroimaging. We demonstrate that this paradigm can identify eye-tracking metrics that differentiate periods of low (Lo) and high (Hi) cognitive demand (evidenced by increased frontal midline theta activity) during driving among a cohort of adolescents ($n$ = 11; mean age =15.1 ± 1.5 yrs). Eye-tracking metrics were compared between a period requiring active brake control to stop at a traffic light (Hi) vs. a period of coasting along a straightaway (Lo). Fixation count ($p<0.02$), mean gaze position ($p<0.01$), and spread of search ($p<0.05$) significantly differed between *braking* (Hi) and *coasting* (Lo). These findings suggest that eye-tracking can be used to detect periods of elevated cognitive control during driving.

## 1. Introduction

Recognition errors – including inadequate surveillance and inattention – are the leading cause of crashes among young drivers (Curry et al., 2011; Lestina & Miller, 1994; McKnight & McKnight, 2003; Seacrist et al., 2021). These errors are largely attributable to limited capacity frontal-lobe cognitive abilities, some of which are still maturing through adolescence and into adulthood (Satterthwaite et al., 2013).

Recent developments in eye-tracking and neuroimaging are paving the way for new approaches to understanding cognitive errors during driving. If we can link fundamental neurocognitive capacities (e.g., frontal lobe cognition important for situational awareness) with driver behaviors that are detectable via in-vehicle eye-tracking, we may be able to detect deficits in drivers' cognitive function in real-time.

Recently, we established a novel paradigm for identifying increased cognitive control during driving by linking magnetoencephalography (MEG)-recorded frequency-specific brain activity to simulated driving performance (E. Walshe et al., 2019; E. A. Walshe et al., 2023). Specifically, we demonstrated increased frontal midline theta (FMT) activity – an established marker of cognitive control over behavior (Callaghan et al., 2017) – when braking in response to a traffic light (requiring top-down cognitive control) relative to rest. However, this study did not include eye-tracking. Eye-tracking metrics may proxy the previously observed FMT cognitive control activity during braking – which may be detectable via driving monitoring systems. To this end, we incorporated eye-tracking into our existing paradigm. This paper aims to demonstrate the utility of this new paradigm to identify eye-tracking metrics associated with periods of increased cognitive control during driving. Building on our prior study, we hypothesize that eye-behavior will vary between periods of high and low cognitive demand.

## 2. Method

This study was approved by the Institutional Review Board at the Children's Hospital of Philadelphia.

### *2.1 Population*

Eleven typically developing adolescents (*Mean Age*: 15.1 ± 1.5 yrs; *Sex*: 4 Female, 7 Male) were recruited. Parental consent and participant assent were obtained. Exclusion criteria were previous diagnosis of autism spectrum disorder, Asperger's syndrome, pervasive developmental delay, other psychiatric disorders, seizure and neurologic disorders, severe claustrophobia, or uncorrectable hearing or vision issues.

### *2.2 Test Procedure*

A detailed description of the MEG, driving simulator hardware, and analysis can be found in Walshe et al. (2023). Briefly, after consent procedures, participants were acclimated to the MEG-compatible driving simulator hardware (Current Designs, Inc., USA) and the simulation (Carnetsoft BV, The Netherlands) via a practice trial for the driving task (described below). Participants then completed the experimental driving task. Following the experimental driving task, participants completed a magnetic resonance imaging brain scan for anatomic localization of MEG-detected brain activity.

### *2.3 Driving Simulation*

The hardware for driving simulation consisted of a MEG-compatible projection screen paired with a MEG-compatible driving hardware including a steering wheel, brake, and accelerator pedals. Eye-tracking was collected at 1000 Hz using a MEG-compatible SR Research EyeLink 1000 (SR Research, Canada) that was fixed to the bottom of the projection screen (see Fig. 1).

Participants drove a repeated-trial prototypical driving scenario (see Fig. 2), beginning at a traffic light intersection, then driving straight to the next intersection, where the traffic turns yellow then red, requiring the participant to brake and stop. No ambient traffic or distractions were present. The driving task was repeated for 20 trials, separated by 9 sec rests.

***Fig 1.*** *Exemplar participant using the driving simulation in the MEG laboratory with MEG-compatible eye-tracker.*

#### *2.3.1 Coasting and Braking Phases:*

To identify eye-tracking metrics related to periods of increased cognitive control, we examined (1) a *Coasting* (Lo) phase: 4s of steady-state driving between acceleration and braking when the gas pedal exhibited the least amount of variation (i.e. the period requiring minimal cognitive control over behavior while still operating the vehicle) and (2) a *Braking* (Hi) phase: a previously established 4s window beginning at the onset of braking in response to an upcoming red light where there is a distinct increase in FMT activity for cognitive control over driving behavior (E. A. Walshe et al., 2023).

#### *2.3.2 MEG Data Collection and Analysis*

Whole-head MEG recordings were conducted using a CTF-Omega 275 channel system (CTF MEG International Services, Coquitlam, B.C.) sampled at 600 Hz. Head position was monitored at 10 Hz. To confirm that the Braking phase represented a period of elevated cognitive control among these younger, newly recruited participants, a frequency-specific differential beamformer-based spatial-filter analysis was used to compare FMT activity (3-9 Hz) (Sakihara et al., 2014; E. A. Walshe et al., 2023) contrasted against a 4 second baseline window (4-8 sec) during the 9 second rest period prior to each trail. FMT activity was visually assessed across all participants.

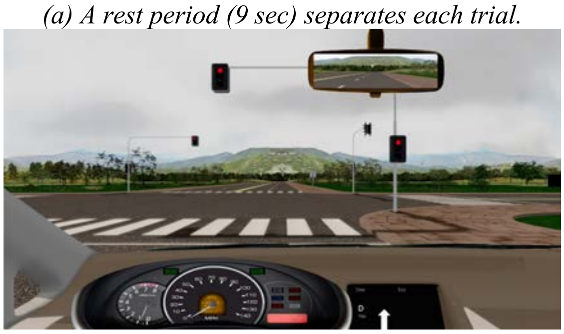
### *2.4 Data Reduction:*

*Coasting* ($n$=48) and *Braking* ($n$=49) phases with more than 20% missing eye-tracking data were excluded from the analysis. The final dataset consisted of 172 *Coasting* and 171 *Braking* phases across all participants. Fixation count, mean fixation duration, mean gaze position, mean saccade amplitude/velocity, and spread of search were computed during the *Coasting* and *Braking* phases for all trials, all participants.

### *2.5 Statistical Analysis*

Paired sample t-tests were used to compare sample means (coasting vs. braking) across the eye-tracking metrics.


*(a) A rest period (9 sec) separates each trial.*



*(b) Participant begins at red light, awaiting green light.*

*(c) Participant proceeds straight to next intersection.*

*(d) The light turns yellow, then red. The participant brakes to come to a stop at the intersection.*

***Fig. 2.*** *Traffic Light Driving Task.*

***Table 1****. Comparison of Eye-Tracking Metrics: Coasting vs. Braking Phases*

| | **Fixation Count** | | **Fixation Duration** | | **Mean Gaze (Horzizontal)** | | **Mean Gaze (Vertical)** | | **Spread of Search (Horizontal)** | | **Spread of Search (Vertical)** | | **Saccade Amplitude** | | **Saccade Velocity** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | # | | sec | | pxls | | pxls | | pxls | | pxls | | pxls | | pxls/sec | |
| | **Coasting** | **Braking** | **Coasting** | **Braking** | **Coasting** | **Braking** | **Coasting** | **Braking** | **Coasting** | **Braking** | **Coasting** | **Braking** | **Coasting** | **Braking** | **Coasting** | **Braking** |
| **Mean** | 11.7 | 9.7 | 0.339 | 0.429 | 839 | 932 | 738 | 652 | 160 | 147 | 174 | 147 | 176 | 162 | 0.057 | 0.055 |
| **SD** | 3.4 | 3.7 | 0.125 | 0.146 | 55 | 72 | 68 | 100 | 33 | 59 | 24 | 38 | 34 | 64 | 0.015 | 0.019 |
| **p-value** | **0.016** | | *0.056* | | **0.004** | | **0.003** | | 0.461 | | **0.049** | | 0.462 | | 0.653 | |

## 3. Results

Participants successfully braked at the traffic light for all trials. Increased FMT activity was observed in all participants during the *Braking* phase.

Mean (±SD) fixation count, fixation duration, mean gaze position, saccade amplitude/velocity, and spread of search are shown in Table 1. Fixation count ($p<0.02$) was significantly lower, vertical spread of search ($p<0.05$) was significantly narrower, and horizontal/vertical mean gaze ($p<0.01$) was significantly different during *Braking* than *Coasting*. A trend ($p=0.06$) toward longer mean fixation duration during *Braking* was also observed.

## 4. Discussion

The goal of this study was to establish the utility of adding eye tracking to a novel paradigm of combined MEG neuroimaging with simulated driving, to allow for identifying eye-tracking metrics associated with periods of increased cognitive control (i.e. braking) during driving. The findings supported our hypothesis that eye-behavior can be used as a proxy for measures of cognitive control during driving. Establishing specific eye-tracking metrics that proxy cognitive control during driving will help inform the development of novel in-vehicle technology that target cognitive errors.

### *4.1 Limitations*

Several limitations warrant discussion. First, the prescribed seating position and lack of head motion permitted in the MEG may have influenced eye-behavior. While the upright seated posture used in the MEG is more naturalistic than the supine position used in prior fMRI studies (Kan et al., 2013), future work should explore whether differences in eye-behavior between coasting and braking are observed in more naturalistic seating postures that allow additional head motion.

Additionally, our driving task represents a relatively simplistic driving scenario. This simplistic task serves as a useful probe of increased cognitive control during driving, without the influence of additional cognitive demands (e.g. ambient traffic). Future work should explore the relationship between eye-behavior and cognitive control in more complex driving tasks.

Finally, participants were younger than typical driving age. Younger adolescents were chosen to limit the influence of driving experience on FMT activity. The younger age of our participants combined with the small sample size may limit the generalizability of these findings. However, these results confirm the utility of this paradigm, which will be used in future work to explore eye-tracking metrics related to periods of cognitive control during driving across varying age and driving experience.

## 5. Conclusions

Our findings suggest that eye-behavior may be a useful proxy of cognitive control during driving. These findings may help optimize emerging driver monitoring systems to detect and mitigate cognitive errors.


## 6. Acknowledgments

The authors would like to thank all the volunteers who participated in this study. The authors would also like to acknowledge John Dell, Rachel Golembski, Erin Verzella, Peter Lam, Shivani Patel, and Na'Keisha Robinson, (RTs) for technical assistance during the study visits. The authors also acknowledge Madeline Griffith, MS for her assistance with set-up and testing of the eye-tracking hardware. We acknowledge in-kind support from Current Designs Inc. who supplied the driving simulation hardware. This study was supported in part by the National Institute of Health NIH-R21 5R21NS118410-02, the Children's Hospital of Philadelphia (CHOP) Foerderer Award, the Annenberg Public Policy Center at the University of Pennsylvania, and the Center for Child Injury Prevention Studies (CChIPS). The content is solely the responsibility of the authors and does not necessarily represent the official views of the listed funding sources.


## References


Callaghan, E., Holland, C., & Kessler, K. (2017). Age-Related Changes in the Ability to Switch between Temporal and Spatial Attention. *Frontiers in Aging Neuroscience*, *9*. https://doi.org/10.3389/fnagi.2017.00028

Curry, A. E., Hafetz, J., Kallan, M. J., Winston, F. K., & Durbin, D. R. (2011). Prevalence of teen driver errors leading to serious motor vehicle crashes. *Accident Analysis & Prevention*, *43*(4), 1285–1290. https://doi.org/10.1016/j.aap.2010.10.019

Kan, K., Schweizer, T. A., Tam, F., & Graham, S. J. (2013). Methodology for functional MRI of simulated driving. *Medical Physics*, *40*(1), 012301. https://doi.org/10.1118/1.4769107

Lestina, D. C., & Miller, T. R. (1994). Characteristics of Crash-Involved Younger Drivers. *Annual Proceedings of the Association for the Advancement of Automotive Medicine*, *38*, 425–437.

McKnight, A. J., & McKnight, A. S. (2003). Young novice drivers: Careless or clueless? *Accident Analysis & Prevention*, *35*(6), 921–925. https://doi.org/10.1016/S0001-4575(02)00100-8
Sakihara, K., Hirata, M., Ebe, K., Kimura, K., Yi Ryu, S., Kono, Y., Muto, N., Yoshioka, M., Yoshimine, T., & Yorifuji, S. (2014). Cerebral oscillatory activity during simulated driving using MEG. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00975
Satterthwaite, T. D., Wolf, D. H., Erus, G., Ruparel, K., Elliott, M. A., Gennatas, E. D., Hopson, R., Jackson, C., Prabhakaran, K., Bilker, W. B., Calkins, M. E., Loughead, J., Smith, A., Roalf, D. R., Hakonarson, H., Verma, R., Davatzikos, C., Gur, R. C., & Gur, R. E. (2013). Functional Maturation of the Executive System during Adolescence. *The Journal of Neuroscience*, *33*(41), 16249–16261. https://doi.org/10.1523/JNEUROSCI.2345-13.2013
Seacrist, T., Maheshwari, J., Sarfare, S., Chingas, G., Thirkill, M., & Loeb, H. S. (2021). In-depth analysis of crash contributing factors and potential ADAS interventions among at-risk drivers using the SHRP 2 naturalistic driving study. *Traffic Injury Prevention*, *22*(sup1), S68–S73. https://doi.org/10.1080/15389588.2021.1979529
Walshe, E. A., Roberts, T. P. L., Ward McIntosh, C., Winston, F. K., Romer, D., & Gaetz, W. (2023). An event-based magnetoencephalography study of simulated driving: Establishing a novel paradigm to probe the dynamic interplay of executive and motor function. *Human Brain Mapping*, hbm.26197. https://doi.org/10.1002/hbm.26197
Walshe, E., Winston, F. K., McIntosh, C. W., Romer, D., Roberts, T., & Gaetz, W. (2019). Magnetoencephalography during Simulated Driving: A New Paradigm for Driver Assessment. *Proceedings of the 10th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design: Driving Assessment 2019*, 182–188. https://doi.org/10.17077/drivingassessment.1693

# Detection of Driving under Influence (DUI) of Alcohol: An Extended Review in Anticipation of Euro NCAP 2026

Jan-Philipp Göbel [1,2,3*], Jan Cedric Mertens [1], Niklas Peuckmann [1,2], Andreas Riener [2]

[1] CARIAD SE, Berliner Ring 2, 38440 Wolfsburg, Germany
[2] Technische Hochschule Ingolstadt, Esplanade 10, 85049 Ingolstadt, Germany
[3] Johannes Kepler University Linz, Altenberger Straße 69, 4040 Linz, Austria
[*] jan-philipp.goebel@cariad.technology

**Abstract: The effective detection of driving under the influence (DUI) of alcohol remains a significant challenge due to the complex interplay of factors affecting driving behavior and the high variability among individuals. This paper presents a comprehensive review of existing methods for detecting alcohol-induced DUI, underscoring the crucial role of human perception, cognition, motor skills, and behavioral changes in impairing driving performance. We discuss traditional DUI detection metrics, including variance in lateral positioning, speed alterations, and changes in acceleration, braking, and reaction times. Additionally, the utilization of non-invasive in-cabin sensors, particularly infrared cameras for monitoring gaze and temperature, is examined. A notable research gap identified is the lack of a unified approach that combines indicators from both driving behavior and in-cabin sensory data for a more accurate and reliable DUI detection system. This gap highlights the need for innovative solutions to enhance.**

## 1. Introduction

Driving under influence (DUI) of alcohol is one of the main factors leading to road accidents, causing 25% of road-deaths in the EU (Antonio Avenoso, 2019) and 31% of road-deaths in the U.S. (Blincoe, Larry [NHTSA], 2023). While nearly all countries in the EU are in line with the EU recommendation of a legal limit for blood alcohol concentration (BAC) for drives of 0.5‰ (European Communities, 2001), it is estimated that 1.5-2% of all kilometres of road traffic in the EU are driven with an illegal BAC (Antonio Avenoso, 2019). Table 1 states the increasing impairment of the driver with raising BAC levels which then leads to an exponentially higher risk of an accident as shown in figure 1.

With the Vison of "A Safer Future of Mobility" and to reduce the personal and social damage caused by alcohol intoxicated drivers, Euro NCAP 2026 is pushing OEMs towards an Impairment Detection, including DUI (Euro NCAP, 2022). Looking ahead to these new standards, we use this review paper to give an extensive overview of the current state of the art regarding the detection of DUI of alcohol.

***Table 1*** *Effects of Alcohol on Driving (NHTSA, 2016)*

| BAC | Effects on Driving |
|---|---|
| .2‰ | Decline in visual functions and multitasking |
| .5‰ | Decline in coordination, ability to track moving objects and response to emergency situations |
| .8‰ | Decline in concentration, short-term memory, perception and speed control |
| 1.0‰ | Reduced ability to maintain lane position and brake appropriately |
| 1.5‰ | Substantial impairment in vehicle control, attention to driving task and perception |

**Outline:** In our literature research, we have focused on two primary areas. Firstly, we examined the immediate effects of alcohol on the human body (section 2). Secondly, we explored methods for identifying intoxicated individuals based on driving behavior and in-cabin sensors (section 3). We conducted searches across major literature databases, including IEEE, Science Direct, PubMed, and Google Scholar, as well as using Google search with various keywords. The title and abstract of each paper were used to determine whether further investigation was warranted. Upon finding an intriguing paper, we proceeded with a reference search procedure.

## 2. Acute Influence of Alcohol on Human Body

Alcohol consumption can have both short-term and long-term effects on the human body. Long-term effects are often considered as the consequence of chronic alcohol consumption over an extended period and include conditions such as liver cirrhosis, cardiovascular diseases, and increased cancer risk (Zakhari, 2006). For DUI however, short term

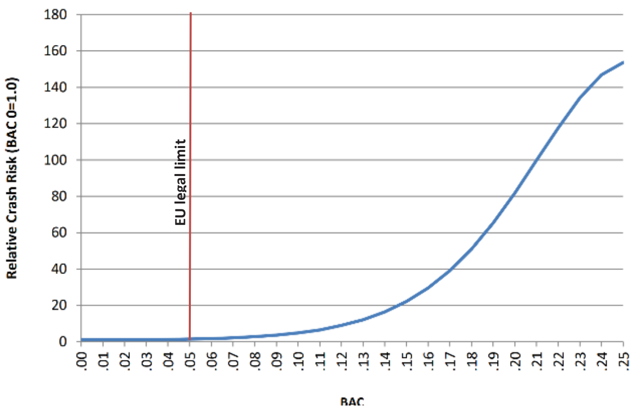



***Fig. 1.*** *Blue: Relative Risk of Crash by Blood Alcohol Concentration (BAC in %) . Red: EU legal limit (Blomberg et al., 2005).*

effects are relevant which occur immediately after alcohol intake and can vary based on factors like individual tolerance, alcohol dosage and overall health (Vonghia et al., 2008). In this section, we'll examine alcohol's general short-term effects on the human body, apart from driving. We reviewed relevant studies and identified seven clusters, summarizing key findings for each. The number in brackets denotes occurrences in the studies. Refer to Appendix A for a complete overview, including participant size, BAC, and influence clusters.

Alcohol is absorbed in the digestive track and transported over the blood to the brain and other organs There it acts as a Central Nervous System (CNS) depressant with sedative effect on the brain activity, causing several impairments (Mitchell, 1985). It can further cause altered behavior and mood, which is the main reason of recreational intake, as well as influence the Cardiovascular system. (See (Mitchell, 1985), (Vonghia et al., 2008) or (Eske & White, 2019) for in-depth explanation of the bio-chemical processes of alcohol in the body). The relevant effects for DUI can be clustered as following:

**Visual Impairment (22):** Reduced ability to efficiently control eye movements (9), change in fixations and saccades (5) and increased pupil dilation (2). In general reduced perception (3) and other impairments (8).
**Cognitive Impairment (15):** Reduced cognitive performance (7), such as attention, planning, recognition. Increased reaction time (2), hallucinations (1) and other impairments (5).
**Behavioral Effects (11):** Change in behavior leading to loss of attention (2), increased aggression (2), euphoria (1), fatigue (1), impulsiveness (1), general social changes (1), emotions (1) and others (2).
**Cardiovascular Effects (8):** Changes in Heart Rate Variability (HRV) (3) and lower indices of cardiac vagal nerve (1). Dilation of blood vessels (1) leading to increased blood flow (1) and higher temperature (2).
**Motoric Impairment (6):** Reduced coordination (4) and balance (2).
**Memory Impairment (3):** Reduced short-term memory capabilities (3).
**Speech Impairment (1):** Slurred speech and reduced articulation (1).

The degree of these short term effects highly depend on the dosage but is also influenced by many other individual factor such as body weight, age and tolerance (Eske & White, 2019). Further short term effects such as Gastrointestinal Effects (Vomiting), Unconscious or Alcohol Poisoning can occur especially at higher BAC levels (Kathleen Davis, 2018). this review however, we considered these as relevant for sudden sickness rather than DUI.

## 3. Methods for DUI Detection

While the previous section focused on the influence of alcohol on the human body we now introduce methods to detect DUI in the vehicle based on the driving behavior or non-invasive in cabin sensors.

### *3.1 Based on Driving Behavior*

The different effects of alcohol on the human body can influence the driving behavior, e.g. not clearly seeing the lane (visual impairment) and not accurately steering the vehicle (motoric impairment) can lead to a lane departure. We identified eight relevant clusters of changed driving behavior used for detecting DUI from the literature review and summarized their key finding, with the number in brackets stating the number of occurrences in the studies. See Appendix B for a full overview of the studies, including Participant Size, BAC, and influence clusters.

**Lane Behavior (21):** Changed variance in lateral position.
**Speed Behavior (18)**: Changed speed variance and average speed.
**Reaction Time (18)**: Changed reaction time.
**Steering Behavior (12):** Changed steering variance, velocity or performance.
**Acceleration Behavior (8)**: Changed Acceleration and speed.
**Braking Behavior (7)**: Changed braking force and reaction time.
**Distance Behavior (5):** Changed distance to leading vehicle and to right lane.
**Others (15):** Changes in Emotions, Attention, Workload Estimation and Perception.

### *3.2 Based on Non-Invasive In-Cabin Sensors*

Some of the influences on alcohol on the human body can be measured with non-invasive in-cabin sensors, such RGB/Near Infrared (NIR)/Far Infrared (FIR) Cameras, Gas sensors and Electrocardiogram (ECG). From the literature review we identified five clusters of approaches taking these sensors input to detect DUI. See Appendix C for a full overview of the studies, including Participant Size, BAC, used method and sensors.

**Gaze Observation (7):** Detecting visual impairments, inattention or a cognitive disconnection from the driving task based on gaze movement.
**Gas Analysis (5):** Passive measurement of BAC in the vehicle air.
**Face Observation (4):** Detecting Emotions from Face Key points.
**Temperature Measurement (3):** Detecting Cardio-vascular Effects based on change in temperature.
**Vital Sign Observations (1):** Detecting Cardio-vascular Effects based on vital signs (e.g. HRV)

## 4. Conclusions

Alcohol intake can have acute impact on the human body causing impairments in perceiving information, (visual), processing them (cognitive) and acting accordingly (motoric). It further leads to a more aggressive behavior and change in vital signs. While some of these effects are hard to measure individually, their combination causes a measurable degradation of driving performance. Key metrics to detect DUI based driving behavior are variance in lateral position and speed, increased acceleration and braking as well as reaction time. In addition, non-invasive sensors in the cabin, especially NIR and FIR cameras, can detect DUI indicators based on gaze and temperature observations.

From the literature review, however, we identified a research gap when combining DUI indicators based on driving behavior and in-cabin sensors. As the DUI detection is due to the high interpersonal differences challenging, all possible indicators should be considered to fulfill NCAP 2026**.**

**References**


Allen, R. W., Jex, H. R., McRuer, D. T., & Dimarco, R. J. (1975). Alcohol Effects on Driving Behavior and Performance in a Car Simulator. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-5*(5), 498–505. https://doi.org/10.1109/TSMC.1975.5408372

Allen, R. W., Parseghian, Z., & Stein, A. C. (1996). A Driving Simulator Study of the Performance Effects of Low Blood Alcohol Concentration. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *40*(18), 943–946. https://doi.org/10.1177/154193129604001817

Altura, B. M., & Altura, B. T. (1984). Alcohol, the cerebral circulation and strokes. *Alcohol (Fayetteville, N.Y.)*, *1*(4), 325–331. https://doi.org/10.1016/0741-8329(84)90056-9

Antonio Avenoso. (2019). *DRINK DRIVING IN THE EU: The fight against drink-driving and alcohol interlocks*. https://etsc.eu/wp-content/uploads/02_-Antonio-Avenoso.pdf

Ariss, T., Fairbairn, C. E., Sayette, M. A., Velia, B. A., Berenbaum, H., & Brown-Schmidt, S. (2023). Where to Look? Alcohol, Affect, and Gaze Behavior During a Virtual Social Interaction. *Clinical Psychological Science*, *11*(2), 239–252. https://doi.org/10.1177/21677026221096449

Arora, S. S., Vatsa, M., Singh, R., & Jain, A. (2012). Iris recognition under alcohol influence: A preliminary study. In *2012 5th IAPR International Conference on Biometrics (ICB)* (pp. 336–341). IEEE. https://doi.org/10.1109/ICB.2012.6199829

Bhango, Z., & van der Haar, D. (2022). A Comparison of Deep Learning Methods for Inebriation Recognition in Humans. In S. Sclaroff, C. Distante, M. Leo, G. M. Farinella, & F. Tombari (Eds.), *Lecture Notes in Computer Science. Image Analysis and Processing – ICIAP 2022* (Vol. 13231, pp. 610–620). Springer International Publishing. https://doi.org/10.1007/978-3-031-06427-2_51

Blincoe, Larry (2023). The Economic and Societal Impact of Motor Vehicle Crashes, 2019 (Revised).

Burns, E. D., Nusbaumer, M. R., & Reiling, D. M. (2003). Think they're drunk? Alcohol servers and the identification of intoxication. *Journal of Drug Education*, *33*(2), 177–186. https://doi.org/10.2190/8U4B-U3VV-WBDN-UKEH

C. A. Naranjo, & K. E. Bremner (1993). Behavioural correlates of alcohol intoxication. *Addiction (Abingdon, England)*, *88*(1), 25–29. https://doi.org/10.1111/j.1360-0443.1993.tb02760.x

Calhoun, V. D., Carvalho, K., Astur, R., & Pearlson, G. D. (2005). Using virtual reality to study alcohol intoxication effects on the neural correlates of simulated driving. *Applied Psychophysiology and Biofeedback*, *30*(3), 285–306. https://doi.org/10.1007/s10484-005-6384-0

Capito, E. S., Lautenbacher, S., & Horn-Hofmann, C. (2017). Acute alcohol effects on facial expressions of emotions in social drinkers: A systematic review. *Psychology Research and Behavior Management*, *10*, 369–385. https://doi.org/10.2147/PRBM.S146918

Chait, L. D., & Perry, J. L. (1994). Acute and residual effects of alcohol and marijuana, alone and in combination, on mood and performance. *Psychopharmacology*, *115*(3), 340–349. https://doi.org/10.1007/BF02245075

Charlton, S. G., & Starkey, N. J. (2015). Driving while drinking: Performance impairments resulting from social drinking. *Accident; Analysis and Prevention*, *74*, 210–217. https://doi.org/10.1016/j.aap.2014.11.001

Childs, E., Roche, D. J. O., King, A. C., & Wit, H. de (2012). Varenicline potentiates alcohol-induced negative subjective responses and offsets impaired eye movements. *Alcoholism, Clinical and Experimental Research*, *36*(5), 906–914. https://doi.org/10.1111/j.1530-0277.2011.01675.x

Christoforou, Z., Karlaftis, M. G., & Yannis, G. (2013). Reaction times of young alcohol-impaired drivers. *Accident; Analysis and Prevention*, *61*, 54–62. https://doi.org/10.1016/j.aap.2012.12.030

Crowdy, K. A., & Marple-Horvat, D. E [D. E.] (2004). Alcohol affects eye movements essential for visually guided stepping. *Alcoholism, Clinical and Experimental Research*, *28*(3), 402–407. https://doi.org/10.1097/01.alc.0000117833.57318.83

Dairi, A., Harrou, F., & Sun, Y. (2022). Efficient Driver Drunk Detection by Sensors: A Manifold Learning-Based Anomaly Detector. *IEEE Access*, *10*, 119001–119012. https://doi.org/10.1109/ACCESS.2022.3221145

Danielle Pacheco, & Anis Rehman. (2023). *Drowsy Driving vs. Drunk Driving: How Similar Are They?* https://www.sleepfoundation.org/drowsy-driving/drowsy-driving-vs-drunk-driving

Dharani, N. P., Ismail, M., & Vidhya, M. (2022). Drunk and Drive Detection System for Safety Driving. In *2022 IEEE International Conference on Current Development in Engineering and Technology (CCET)* (pp. 1–5). IEEE. https://doi.org/10.1109/CCET56606.2022.10080037

do Canto-Pereira, L. H. M., P A David, I. de, Machado-Pinheiro, W., & Ranvaud, R. D. (2007). Effects of acute alcohol intoxication on visuospatial attention. *Human & Experimental Toxicology*, *26*(4), 311–319. https://doi.org/10.1177/0960327106070490

Eckstein, M. K., Guerra-Carrillo, B., Miller Singley, A. T., & Bunge, S. A. (2017). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive*

*Neuroscience*, *25*, 69–91. https://doi.org/10.1016/j.dcn.2016.11.001
Eske, J., & White, C. (2019). *What to know about alcohol intoxication*. https://www.medicalnewstoday.com/articles/327202
Euro NCAP (2022). Euro NCAP Vision 2030: a safer future for mobility. https://cdn.euroncap.com/media/74468/euro-ncap-roadmap-vision-2030.pdf
European Communities. (2001). *COMMISSION RECOMMENDATION: on the maximum permitted blood alcohol content (BAC) for drivers of motorised vehicles*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32001Y0214(01)&from=EN
Fillmore, M. T., Blackburn, J. S., & Harrison, E. L. R. (2008). Acute disinhibiting effects of alcohol as a factor in risky driving behavior. *Drug and Alcohol Dependence*, *95*(1-2), 97–106. https://doi.org/10.1016/j.drugalcdep.2007.12.018
Garg, D., Kumar Srivastava, A., Paliwal, D., Shekhar, S., & Singh Chauhan, A. (2020). Alcohol Detection System in Vehicle Using Arduino. *Journal of Xi' an Shiyou University, Natural Science Edition*. https://www.semanticscholar.org/paper/Alcohol-Detection-System-in-Vehicle-Using-Arduino-Garg-Srivastava/b3f6ec0fafdafb22410d553ccf7e396fbb574dbf#related-papers
Garrisson, H., Scholey, A., Verster, J. C., Shiferaw, B., & Benson, S. (2022). Effects of alcohol intoxication on driving performance, confidence in driving ability, and psychomotor function: A randomized, double-blind, placebo-controlled study. *Psychopharmacology*, *239*(12), 3893–3902. https://doi.org/10.1007/s00213-022-06260-z
Georgia Koukiou. (2017). Intoxication Identification Using Thermal Imaging: 8. In Gholamreza Anbarjafari & Sergio Escalera (Eds.), *Human-Robot Interaction.* IntechOpen. https://doi.org/10.5772/intechopen.72128
Gundersen, H., Grüner, R., Specht, K., & Hugdahl, K. (2008). The effects of alcohol intoxication on neuronal activation at different levels of cognitive load. *The Open Neuroimaging Journal*, *2*, 65–72. https://doi.org/10.2174/1874440000802010065
Hack, M. A., Choi, S. J., Vijayapalan, P., Davies, R. J., & Stradling, J. R. (2001). Comparison of the effects of sleep deprivation, alcohol and obstructive sleep apnoea (OSA) on simulated steering performance. *Respiratory Medicine*, *95*(7), 594–601. https://doi.org/10.1053/rmed.2001.1109
Hafstrom, A., Patel, M., Modig, F., Magnusson, M., & Fransson, P. A. (2014). Acute alcohol intoxication impairs segmental body alignment in upright standing. *Journal of Vestibular Research : Equilibrium & Orientation*, *24*(4), 297–304. https://doi.org/10.3233/VES-140513
Harrison, E. L. R., & Fillmore, M. T. (2005). Are bad drivers more impaired by alcohol? Sober driving precision predicts impairment from alcohol in a simulated driving task. *Accident; Analysis and Prevention*, *37*(5), 882–889. https://doi.org/10.1016/j.aap.2005.04.005
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
Harvey, A. J. (2016). When alcohol narrows the field of focal attention. *Quarterly Journal of Experimental Psychology (2006)*, *69*(4), 669–677. https://doi.org/10.1080/17470218.2015.1040803
Heishman, S. J., Arasteh, K., & Stitzer, M. L. (1997). Comparative effects of alcohol and marijuana on mood, memory, and performance. *Pharmacology, Biochemistry, and Behavior*, *58*(1), 93–101. https://doi.org/10.1016/S0091-3057(96)00456-X
Helland, A., Jenssen, G. D., Lervåg, L.-E., Westin, A. A., Moen, T., Sakshaug, K., Lydersen, S., Mørland, J., & Slørdal, L. (2013). Comparison of driving simulator performance with real driving after alcohol intake: A randomised, single blind, placebo-controlled, cross-over trial. *Accident; Analysis and Prevention*, *53*, 9–16. https://doi.org/10.1016/j.aap.2012.12.042
Helland, A., Lydersen, S., Lervåg, L.-E., Jenssen, G. D., Mørland, J., & Slørdal, L. (2016). Driving simulator sickness: Impact on driving performance, influence of blood alcohol concentration, and effect of repeated simulator exposures. *Accident; Analysis and Prevention*, *94*, 180–187. https://doi.org/10.1016/j.aap.2016.05.008
Hermosilla, G., Verdugo, J. L., Farias, G., Vera, E., Pizarro, F., & Machuca, M. (2018). Face Recognition and Drunk Classification Using Infrared Face Images. *Journal of Sensors*, *2018*, 1–8. https://doi.org/10.1155/2018/5813514
Islam, M. H., Khandoker, A. A., Safi Sami, T. M., Talukder, T. I., Rahman, M. I., & Sarkar, P. K. (2021). Car Accident Prevention And Health Monitoring System For Drivers. In *2021 IEEE Region 10 Symposium (TENSYMP)* (pp. 1–6). IEEE. https://doi.org/10.1109/TENSYMP52854.2021.9551006
Jolkovsky, E. L., Fernandez-Penny, F. E., Alexis, M., Benson, L. N., Wang, B. H., & Abella, B. S. (2022). Impact of acute intoxication on quantitative pupillometry assessment in the emergency department. *Journal of the American College of Emergency Physicians Open*, *3*(5), e12825. https://doi.org/10.1002/emp2.12825
Kathleen Davis. (2018). *What effects does alcohol have on health?* https://www.medicalnewstoday.com/articles/305062
Koukiou, G [G.], & Anastassopoulos, V [V.] (2011). Facial blood vessels activity in drunk persons using thermal infrared. In *4th International Conference on Imaging for Crime Detection and Prevention 2011 (ICDP 2011)* (P11-P11). IET. https://doi.org/10.1049/ic.2011.0108
Koukiou, G [Georgia], & Anastassopoulos, V [Vassilis]. (2013). Eye temperature distribution in drunk persons using thermal imagery. In *GI Edition*

*Proceedings: Vol. 212. Biosig 2013: Proceedings of the 12th international conference of the Biometrics Special Interest Group ; 04. - 06. September 2013 in Darmstadt, Germany* (pp. 1–8). Ges. für Informatik.

Kumar, A [Ankit], Kumar, A [Ajay], Singh, M., Kumar, P., & Bijalwan, A. (2022). An Optimized Approach Using Transfer Learning to Detect Drunk Driving. *Scientific Programming*, *2022*, 1–12. https://doi.org/10.1155/2022/8775607

Lee, K. H., Baek, K. H., Choi, S. B., Jeong, N. T., Moon, H. U., Lee, E. S., Kim, H. M., & Suh, M. W. (2019). Development of Three Driver State Detection Models from Driving Information Using Vehicle Simulator; Normal, Drowsy and Drunk Driving. *International Journal of Automotive Technology*, *20*(6), 1205–1219. https://doi.org/10.1007/s12239-019-0113-9

Li, Y. C., Sze, N. N., Wong, S. C., Yan, W., Tsui, K. L., & So, F. L. (2016). A simulation study of the effects of alcohol on driving performance in a Chinese population. *Accident; Analysis and Prevention*, *95*(Pt B), 334–342. https://doi.org/10.1016/j.aap.2016.01.010

Liguori, A., D'Agostino, R. B., Dworkin, S. I., Edwards, D., & Robinson, J. H. (1999). Alcohol Effects on Mood, Equilibrium, and Simulated Driving. *Alcoholism: Clinical and Experimental Research*, *23*(5), 815–821. https://doi.org/10.1111/j.1530-0277.1999.tb04188.x

Ljungblad, J., Hök, B., Allalou, A., & Pettersson, H. (2017). Passive in-vehicle driver breath alcohol detection using advanced sensor signal acquisition and fusion. *Traffic Injury Prevention*, *18*(sup1), S31-S36. https://doi.org/10.1080/15389588.2017.1312688

Lobato-Rincón, L. L., Cabanillas Campos, M. C., Navarro-Valls, J. J., Bonnin-Arias, C., Chamorro, E., & Sánchez-Ramos Roda, C. (2013). Utilidad de la pupilometría dinámica en el control de alcoholemia de los conductores [Utility of dynamic pupillometry in alcohol testing on drivers]. *Adicciones*, *25*(2), 137–145.

Ma, J [Jin], Ma, R., Liu, X., Wen, Z., Li, X., Wang, T., Han, W., Hu, W., & Zhang, Z. (2011). Correlation between heart rate variability and pupillary reflex in healthy adult subjects under the influence of alcohol. Advance online publication. https://doi.org/10.3969/j.issn.1673-5374.2011.21.008

Makowski, S., Prasse, P., Jager, L. A., & Scheffer, T. (2022). Oculomotoric Biometric Identification under the Influence of Alcohol and Fatigue. In *2022 IEEE International Joint Conference on Biometrics (IJCB)* (pp. 1–9). IEEE. https://doi.org/10.1109/IJCB54206.2022.10007970

Maria J. Perez Carrasco, Sr., L. L. L.-R., Jose G. Ramirez-Mercado, Maria del Carmen Cabanillas-Campos, Gonzalo Alvarez-Rementeria, Antonio Langa-Moraga, & Celia Sanchez-Ramos Sr. (2011). ffects Of Alcohol On The Mesopic Pupil Reflex Produced In Response To Coloured Flash Stimuli. *Carrasco, M.J., Lobato-Rincón, L.L., Ramirez-Mercado, J.G., Cabanillas-Campos, M.D., Alvarez-Rementeria, G., Langa‑Moraga, A., & Sánchez‑ramos, C. (2011). Effects of Alcohol on the Mesopic Pupil Reflex Produced in Response to Coloured Flash Stimuli. Investigative Ophthalmology & Visual Science, 52*, 4909.

Marple-Horvat, D. E [Dilwyn E.], Cooper, H. L., Gilbey, S. L., Watson, J. C., Mehta, N., Kaur-Mann, D., Wilson, M., & Keil, D. (2008). Alcohol badly affects eye movements linked to steering, providing for automatic in-car detection of drink driving. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *33*(4), 849–858. https://doi.org/10.1038/sj.npp.1301458

Mehta, V., Yadav, D. P., Katta, S. S., & Dhall, A. (2018). *DIF : Dataset of Perceived Intoxicated Faces for Drunk Person Identification.* https://doi.org/10.48550/arXiv.1805.10030

Mitchell, M. C. (1985). Alcohol-induced impairment of central nervous system function: Behavioral skills involved in driving. *Journal of Studies on Alcohol. Supplement*, *10*, 109–116. https://doi.org/10.15288/jsas.1985.s10.109

National Highway Traffic Safety Administration. (2016). *The ABCs of BAC: A Guide to Understanding Blood Alcohol Concentration and Alcohol Impairment*. https://www.nhtsa.gov/sites/nhtsa.gov/files/809844-theabcsofbac.pdf

Patel, M., Modig, F., Magnusson, M., & Fransson, P. A. (2010). Alcohol intoxication at 0.06 and 0.10% blood alcohol concentration changes segmental body movement coordination. *Experimental Brain Research*, *202*(2), 431–443. https://doi.org/10.1007/s00221-009-2150-5

Penton-Voak, I. S., Cooper, R. M., Roberts, R. E., Attwood, A. S., & Munafò, M. R. (2012). Effects of acute alcohol consumption on the perception of eye gaze direction. *Journal of Psychopharmacology (Oxford, England)*, *26*(2), 254–261. https://doi.org/10.1177/0269881110385599

Persson, L. O., Sjöberg, L., & Svensson, E. (1980). Mood effects of alcohol. *Psychopharmacology*, *68*(3), 295–299. https://doi.org/10.1007/BF00428119

Peterson, J. B., Rothfleisch, J., Zelazo, P. D., & Pihl, R. O. (1990). Acute alcohol intoxication and cognitive functioning. *Journal of Studies on Alcohol*, *51*(2), 114–122. https://doi.org/10.15288/jsa.1990.51.114

Quintana, D. S., Guastella, A. J., McGregor, I. S., Hickie, I. B., & Kemp, A. H. (2013). Moderate alcohol intake is related to increased heart rate variability in young adults: Implications for health and well-being. *Psychophysiology*, *50*(12), 1202–1208. https://doi.org/10.1111/psyp.12134

Roche, D. J. O., & King, A. C. (2010). Alcohol impairment of saccadic and smooth pursuit eye movements: Impact of risk factors for alcohol dependence. *Psychopharmacology*, *212*(1), 33–44. https://doi.org/10.1007/s00213-010-1906-8

Romanowicz, M., Schmidt, J. E., Bostwick, J. M., Mrazek, D. A., & Karpyak, V. M. (2011). Changes in heart rate variability associated with acute alcohol consumption: Current knowledge and implications for practice and research. *Alcoholism: Clinical and Experimental Research*, *35*(6), 1092–1105. https://doi.org/10.1111/j.1530-0277.2011.01442.x

Rose, A. K., Brown, K., MacKillop, J., Field, M., & Hogarth, L. (2018). Alcohol devaluation has dissociable effects on distinct components of alcohol behaviour. *Psychopharmacology*, *235*(4), 1233–1244. https://doi.org/10.1007/s00213-018-4839-2

Ryan, J. M., & Howes, L. G. (2002). Relations between alcohol consumption, heart rate, and heart rate variability in men. *Heart (British Cardiac Society)*, *88*(6), 641–642. https://doi.org/10.1136/heart.88.6.641

Schweizer, T. A., & Vogel-Sprott, M. (2008). Alcohol-impaired speed and accuracy of cognitive functions: A review of acute tolerance and recovery of cognitive performance. *Experimental and Clinical Psychopharmacology*, *16*(3), 240–250. https://doi.org/10.1037/1064-1297.16.3.240

Shiferaw, B. A., Crewther, D. P., & Downey, L. A. (2019). Gaze entropy measures detect alcohol-induced driver impairment. *Drug and Alcohol Dependence*, *204*, 107519. https://doi.org/10.1016/j.drugalcdep.2019.06.021

Shirazi, M. M., & Rad, A. B. (2014). Detection of Intoxicated Drivers Using Online System Identification of Steering Behavior. *IEEE Transactions on Intelligent Transportation Systems*, *15*(4), 1738–1747. https://doi.org/10.1109/TITS.2014.2307891

Silva, J. B. S., Cristino, E. D., Almeida, N. L. de, Medeiros, P. C. B. de, & Santos, N. A. D. (2017). Effects of acute alcohol ingestion on eye movements and cognition: A double-blind, placebo-controlled study. *PloS One*, *12*(10), e0186061. https://doi.org/10.1371/journal.pone.0186061

Silvia Makowski, Annika Bätz, Paul Prasse, Lena A. Jäger, & Tobias Scheffer (2023). Detection of Alcohol Inebriation from Eye Movements. https://www.uni-potsdam.de/fileadmin/projects/cs-ml/media/KES2023.pdf

Strayer, D. L., Drews, F. A., & Crouch, D. J. (2006). A comparison of the cell phone driver and the drunk driver. *Human Factors*, *48*(2), 381–391. https://doi.org/10.1518/001872006777724471

Tapia, J. E., Droguett, E. L., Valenzuela, A., Benalcazar, D. P., Causa, L., & Busch, C. (2021). Semantic Segmentation of Periocular Near-Infra-Red Eye Images Under Alcohol Effects. *IEEE Access*, *9*, 109732–109744. https://doi.org/10.1109/ACCESS.2021.3101256

Tarter, R. E., Jones, B. M., Simpson, C. D., & Vega, A. (1971). Effects of task complexity and practice on performance during acute alcohol intoxication. *Perceptual and Motor Skills*, *33*(1), 307–318. https://doi.org/10.2466/pms.1971.33.1.307

Volkow, N. D., Mullani, N., Gould, L., Adler, S. S., Guynn, R. W., Overall, J. E., & Dewey, S. (1988). Effects of acute alcohol intoxication on cerebral blood flow measured with PET. *Psychiatry Research*, *24*(2), 201–209. https://doi.org/10.1016/0165-1781(88)90063-7

Vonghia, L., Leggio, L., Ferrulli, A., Bertini, M., Gasbarrini, G., & Addolorato, G. (2008). Acute alcohol intoxication. *European Journal of Internal Medicine*, *19*(8), 561–567. https://doi.org/10.1016/j.ejim.2007.06.033

Wang, H., Wang, Z., Zhou, Y., & Tzvetanov, T. (2021). Moderate Alcohol Intake Changes Visual Perception by Enhancing V1 Inhibitory Surround Interactions. *Frontiers in Neuroscience*, *15*, Article 682229. https://doi.org/10.3389/fnins.2021.682229

Watten, R. G., & Lie, I. (1997). The effects of alcohol on eye movements during reading. *Alcohol and Alcoholism (Oxford, Oxfordshire)*, *32*(3), 275–280. https://doi.org/10.1093/oxfordjournals.alcalc.a008267

Wester, A. E., Verster, J. C., Volkerts, E. R., Böcker, K. B. E., & Kenemans, J. L. (2010). Effects of alcohol on attention orienting and dual-task performance during simulated driving: An event-related potential study. *Journal of Psychopharmacology (Oxford, England)*, *24*(9), 1333–1348. https://doi.org/10.1177/0269881109348168

Wu, C. K., Tsang, K. F., Chi, H. R., & Hung, F. H. (2016). A Precise Drunk Driving Detection Using Weighted Kernel Based on Electrocardiogram. *Sensors (Basel, Switzerland)*, *16*(5). https://doi.org/10.3390/s16050659

Wu, Z., Feng, C., Zhang, X [Xinting], & Chen, G. (2011). Effects of alcohol intoxication on simulated driving performance. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (pp. 1852–1857). IEEE. https://doi.org/10.1109/ITSC.2011.6083124

Yadav, A. K., & Velaga, N. R. (2019). Effect of alcohol use on accelerating and braking behaviors of drivers. *Traffic Injury Prevention*, *20*(4), 353–358. https://doi.org/10.1080/15389588.2019.1587167

Yoda, T., Crawshaw, L. I., Nakamura, M., Saito, K., Konishi, A., Nagashima, K., Uchida, S., & Kanosue, K. (2005). Effects of alcohol on thermoregulation during mild heat exposure in humans. *Alcohol (Fayetteville, N.Y.)*, *36*(3), 195–200. https://doi.org/10.1016/j.alcohol.2005.09.002

Zakhari, S. (2006). Overview: How Is Alcohol Metabolized by the Body? *Alcohol Research & Health*, *29*(4), 245–254.

Zhang, P., Guo, Y., Qiao, Y., Yan, N., Zhang, Y., Ren, W., Zhang, S., & Di Wu (2022). Acute Alcohol Intake Affects Internal Additive Noise and the Perceptual Template in Visual Perception. *Frontiers in Neuroscience*, *16*, 873671. https://doi.org/10.3389/fnins.2022.873671

Zhang, X [Xingjian], Zhao, X., Du, H., Ma, J [Jianming], & Rong, J. (2014). Effect of different breath alcohol concentrations on driving performance in horizontal curves. *Accident; Analysis and Prevention*, *72*, 401–410. https://doi.org/10.1016/j.aap.2014.07.032

Zube, K., Daldrup, T., Lau, M., Maatz, R., Tank, A., Steiner, I., Schwender, H., & Hartung, B. (2022). Escooter driving under the acute influence of alcohol-a real-driving fitness study. *International Journal of Legal Medicine*, *136*(5), 1281–1290. https://doi.org/10.1007/s00414-022-02792-3

## Appendix A

***Table 2** Overview of Studies investigation the short-term influence of alcohol on the human body*
*(N: Sample size, BAC: Different Blood Alcohol Concentration in ‰ investigated, .0P: Placebo, ●: relevant Cluster)*

| Source | N | BAC [‰] | Visual | Motoric | Speech | Memory | Cognitive | Behaviour | Cardiovascular |
|---|---|---|---|---|---|---|---|---|---|
| (Silva et al., 2017) | 20 | .0P/.8 | ● | ○ | ○ | ○ | ● | ○ | ○ |
| (P. Zhang et al., 2022) | 9 | ~.0/.8 | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| (Eckstein et al., 2017) | | | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| (Wang et al., 2021) | 33 | ~.0/.6 | ● | ○ | ○ | ○ | ● | ○ | ○ |
| (Ariss et al., 2023) | 246 | .0/.8 | ● | ○ | ○ | ○ | ○ | ● | ○ |
| (Roche & King, 2010) | 138 | .0/.4/.8 | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| (Watten & Lie, 1997) | 18 | .0/.5/1.0 | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| (Helland et al., 2016) | 20 | .0/.5/.9 | ○ | ○ | ○ | ○ | ○ | ● | ○ |
| (Childs et al., 2012) | 13 | ~.0/.4/.8 | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| (Harvey, 2016) | 32 | .0P/.95 | ● | ○ | ○ | ○ | ○ | ● | ○ |
| (Rose et al., 2018) | 62 | | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| (Heishman et al., 1997) | 5 | .0P/.25/.5/1.0 | ○ | ○ | ○ | ● | ● | ○ | ○ |
| (Crowdy & Marple-Horvat, 2004) | 6 | ~.5 | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| (Burns et al., 2003) | 822 | | ○ | ○ | ○ | ○ | ○ | ● | ○ |
| (Georgia Koukiou & Anastassopoulos, 2013) | 40 | ~.5 | ○ | ○ | ○ | ○ | ○ | ○ | ● |
| (Altura & Altura, 1984) | | >.25 | ○ | ○ | ○ | ○ | ● | ○ | ○ |
| (Schweizer & Vogel-Sprott, 2008) | | | ○ | ○ | ○ | ○ | ● | ○ | ○ |
| (Penton-Voak et al., 2012) | 30 | .0/.2/.4 | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| (Persson et al., 1980) | 11 | .0/.41/.63/.85 | ○ | ○ | ○ | ○ | ○ | ● | ○ |
| (Chait & Perry, 1994) | 14 | ~.0/.6 | ○ | ○ | ○ | ○ | ○ | ● | ○ |
| (Ma et al., 2011) | 16 | .0/.5/.8 | ● | ○ | ○ | ○ | ○ | ● | ● |
| (do Canto-Pereira et al., 2007) | 24 | .0/.8 | ● | ○ | ○ | ○ | ● | ○ | ○ |
| (G. Koukiou & Anastassopoulos, 2011) | 20 | ~.0/.75 | ○ | ○ | ○ | ○ | ○ | ○ | ● |
| (Jolkovsky et al., 2022) | 407 | | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| (C. A. Naranjo & K. E. Bremner, 1993) | | | ○ | ○ | ○ | ○ | ○ | ● | ○ |
| (Liguori et al., 1999) | 18 | .0/.5/.8 | ○ | ○ | ○ | ○ | ● | ○ | ○ |
| (Peterson et al., 1990) | 72 | ~.0/.13/.66 | ○ | ● | ● | ● | ● | ○ | ○ |
| (Mitchell, 1985) | | | ● | ● | ○ | ○ | ● | ○ | ○ |
| (Tarter et al., 1971) | 26 | ~.8 | ● | ● | ○ | ○ | ● | ○ | ○ |
| (Volkow et al., 1988) | 13 | .0/.5/1.0 | ○ | ○ | ○ | ○ | ○ | ○ | ● |
| (Shiferaw et al., 2019) | 22 | ~.6 | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| (Calhoun et al., 2005) | 19 | .0/.4/.8 | ● | ○ | ○ | ○ | ● | ○ | ○ |
| (Zube et al., 2022) | 57 | ~.3/.6 | ○ | ● | ○ | ○ | ○ | ○ | ○ |
| (Maria J. Perez Carrasco et al., 2011) | 2 | | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| (Lobato-Rincón et al., 2013) | 19 | ~.5 | ● | ○ | ○ | ○ | ○ | ○ | ○ |
| (Danielle Pacheco & Anis Rehman, 2023) | | .5/.8/1.0 | ○ | ○ | ○ | ○ | ● | ● | ○ |
| (Marple-Horvat et al., 2008) | 10 | .7 | ● | ○ | ○ | ○ | ● | ○ | ○ |
| (Patel et al., 2010) | 25 | .6/1.0 | ● | ● | ○ | ○ | ○ | ○ | ○ |
| (Christoforou et al., 2013) | 49 | | ○ | ○ | ○ | ○ | ● | ○ | ○ |
| (Ryan & Howes, 2002) | 39 | | ○ | ○ | ○ | ○ | ○ | ○ | ● |
| (Romanowicz et al., 2011) | | | ○ | ○ | ○ | ○ | ○ | ○ | ● |
| (Gundersen et al., 2008) | 25 | .8 | ○ | ○ | ○ | ● | ● | ○ | ○ |
| (Quintana et al., 2013) | 47 | | ○ | ○ | ○ | ○ | ○ | ○ | ● |
| (Hafstrom et al., 2014) | 25 | .6/1.0 | ○ | ● | ○ | ○ | ○ | ○ | ○ |
| (Capito et al., 2017) | | | ○ | ○ | ○ | ○ | ○ | ● | ○ |
| (Yoda et al., 2005) | 8 | ~.36 | ○ | ○ | ○ | ○ | ○ | ○ | ● |

## Appendix B

***Table 3*** *Overview of Studies investigating DUI detection based on driving behavior*
*(N: Sample size, BAC: Different Blood Alcohol Concentration in ‰ investigated, .0P: Placebo, ●: relevant Cluster, FVG: Use of BAC equivalent Fatal Vision Googles)*

| Study | N | BAC [‰] | Acceleration | Speed | Braking | Lane | Steering | Distance | Reaction | Others |
|---|---|---|---|---|---|---|---|---|---|---|
| (Strayer et al., 2006) | 40 | .8 | ○ | ● | ● | ○ | ○ | ● | ● | ○ |
| (Allen et al., 1996) | 33 | .55 | ● | ○ | ○ | ● | ○ | ○ | ● | ○ |
| (Li et al., 2016) | 52 | | ○ | ● | ○ | ● | ○ | ○ | ● | ○ |
| (Fillmore et al., 2008) | 14 | .8 | ● | ● | ○ | ● | ● | ○ | ○ | ● |
| (Harrison & Fillmore, 2011) | 40 | .65 | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ |
| (Marple-Horvat et al., 2008) | 10 | .7 | ○ | ● | ○ | ○ | ● | ○ | ● | ● |
| (Allen et al., 1975) | 18 | .0/.6/1.1 | ○ | ○ | ○ | ● | ○ | ○ | ● | ● |
| (Harrison & Fillmore, 2005) | 28 | .65 | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ |
| (Helland et al., 2013) | 20 | .0/.5/.9 | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ |
| (Hack et al., 2001) | 12 | .71 | ○ | ○ | ○ | ● | ● | ○ | ● | ○ |
| (Shirazi & Rad, 2014) | | .17-.2 (FVG) | ○ | ● | ○ | ● | ● | ○ | ○ | ● |
| (Lee et al., 2019) | 12 | ≥ .5 | ● | ○ | ○ | ● | ● | ○ | ● | ○ |
| (Charlton & Starkey, 2015) | 44 | | ○ | ● | ○ | ○ | ● | ○ | ○ | ○ |
| (Yadav & Velaga, 2019) | 82 | .0/.3/.5/.8 | ● | ○ | ● | ○ | ○ | ○ | ○ | ● |
| (X. Zhang et al., 2014) | 25 | .0/.3/.6/.9 | ○ | ● | ○ | ○ | ● | ● | ○ | ○ |
| (Garrisson et al., 2022) | 17 | .5/.8 | ○ | ● | ○ | ● | ○ | ○ | ● | ● |
| (Z. Wu et al., 2011) | 13 | .0/.2/.5/.8 | ● | ○ | ● | ○ | ○ | ○ | ● | ● |
| (Wester et al., 2010) | 32 | .0/.2/.5/.8/1.0 | ○ | ○ | ○ | ○ | ○ | ○ | ● | ○ |
| (Mets et al., 2011) | 27 | .5/.8/1.1 | ○ | ● | ○ | ● | ○ | ○ | ○ | ● |
| (Christoforou et al., 2012) | 49 | | ○ | ● | ○ | ● | ● | ● | ● | ● |
| (Z. Li et al., 2019) | 25 | .0/.3/.6/.9 | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ |
| (Y.-C. Liu & Ho, 2010) | 8 | .0/.5/.8/1.0 | ○ | ● | ○ | ○ | ○ | ● | ● | ● |
| (Ou et al., 2010) | | | ● | ● | ● | ○ | ● | ○ | ○ | ○ |
| (Čulík et al., 2022) | 30 | | ○ | ○ | ○ | ○ | ○ | ○ | ● | ○ |
| (Leung & Starmer, 2005) | 32 | .0/.7 | ○ | ○ | ○ | ○ | ○ | ○ | ● | ○ |
| (Zhao et al., 2011) | 24 | ≥ .5 | ● | ● | ● | ○ | ● | ○ | ○ | ○ |
| (Ramaekers et al., 2000) | 18 | .4 | ○ | ○ | ○ | ● | ○ | ○ | ● | ● |
| (Yadav & Velaga, 2021) | 82 | .0/.3/.5/.8 | ○ | ○ | ○ | ● | ● | ○ | ○ | ○ |
| (Yadav & Velaga, 2019b) | 79 | .0/.3/.5/.8 | ○ | ○ | ○ | ○ | ○ | ○ | ● | ○ |
| (Christoforou et al., 2013) | 49 | | ○ | ○ | ○ | ○ | ○ | ○ | ● | ○ |
| (Weafer et al., 2008) | 24 | .0/.45/.65 | ○ | ● | ○ | ● | ○ | ○ | ○ | ○ |
| (Zhao et al., 2014) | 25 | .0/.3/.6/.9 | ○ | ● | ○ | ● | ○ | ○ | ○ | ○ |
| (Y. C. Liu & Ho, 2007) | 8 | .0/.25/.4/.5 | ○ | ● | ○ | ● | ○ | ○ | ○ | ● |
| (Ján Vrábel & Zuzana Majerová, 2013) | | | ○ | ○ | ● | ● | ○ | ○ | ● | ● |
| (McCartney et al., 2017) | 22 | 0.6 (FGV) | ○ | ● | ○ | ● | ○ | ● | ● | ● |
| (Vollrath & Fischer, 2017) | 48/63 | 0.5 | ○ | ● | ● | ○ | ○ | ○ | ● | ● |

**Appendix C**

***Table 4*** *Overview of Studies investigating DUI detection with Non-Invasive In-Cabin Sensors*
*(N: Sample size, BAC: Different Blood Alcohol Concentration in ‰ investigated, ●: relevant Cluster, d: Data points)*

| Study | N | BAC [‰] | Method | | | | | Sensor | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Gaze | Face | Temperature | Gas | Vital Signs | NIR Camera | RGB Camera | FIR Camera | Gas Sensor | ECG | Driving Data |
| (Silvia Makowski et al., 2023) | 44 | .0/.8 | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ○ |
| (Watten & Lie, 1997) | 26 | .0/1.0 | ● | ○ | ○ | ○ | ○ | ● | ? | ○ | ○ | ○ | ○ |
| (Marple-Horvat et al., 2008) | 10 | .0/<.8 | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ● |
| (Makowski et al., 2022) | 66 | .0/.5 | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ○ |
| (Tapia et al., 2021) | 266 | .0/? | ● | ○ | ○ | ○ | ○ | ● | ? | ○ | ○ | ○ | ○ |
| (Arora et al., 2012) | 55 | .0/1.0 | ● | ○ | ○ | ○ | ○ | ● | ? | ○ | ○ | ○ | ○ |
| (Kumar et al., 2022) | | | ○ | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ |
| (Georgia Koukiou, 2017) | 41 | .0/~0.5 | ○ | ○ | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ |
| (Dharani et al., 2022) | | | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ |
| (Garg et al., 2020) | | | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ |
| (Islam et al., 2021) | | | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ |
| (C. K. Wu et al., 2016) | 50 | .0/>.2 | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ | ● | ○ |
| (Dairi et al., 2022) | 390d | | ● | ○ | ● | ● | ○ | ○ | ● | ● | ● | ○ | ○ |
| (Ljungblad et al., 2017) | 10 | .0/.3 | ○ | ● | ○ | ● | ○ | ● | ○ | ● | ● | ○ | ○ |
| (Hermosilla et al., 2018) | 46 | .0/? | ○ | ○ | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ |
| (Mehta et al., 2018) | | | ○ | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ |
| (Bhango & van der Haar, 2022) | 5008d | | ○ | ● | ○ | ○ | ○ | ○ | ● | ○ | ○ | ○ | ○ |

# Angry drivers: A simulator study on investigating on/off-road anger

Jordan Maillant [12*], Stéphanie Dabic [1], Christophe Jallais [2]

[1] Valeo Switches & Smart Controls Product Line , Rue Jules VERNE - Vetraz Monthoux
BP 509 - 74106 Annemasse Cedex, FRANCE
[2] Univ Gustave Eiffel, Univ Lyon, TS2-LESCOT, F-69675 Lyon, France
[*]jordan.maillant@valeo.com

**Abstract:**

**This study aimed to assess whether the source of the driver's anger (whether related to the driving environment or not) is associated with different ocular and physiological responses. 37 volunteers took part in this driving simulator study. They were divided into 4 groups: neutral thoughts and neutral driving events (NN); neutral thoughts and angry driving events (NA); angry thoughts and neutral driving events (AN); angry thoughts and angry driving events (AA). Subjective assessments as well as cardiac, respiratory, electrodermal and ocular activities were collected. Subjective data revealed the effectiveness of the protocol in inducing anger either from thoughts and driving events but even more by a combination of both. Preliminary analysis concerning ocular behaviors showed that participants in the control group (NN) had narrower vertical gaze explorations, less fixations and saccades and longer fixation duration than anger-induced participants. Nevertheless, the oculometric data do not allow us to make distinctions between the two sources of anger. These results, and the analysis to follow of the other indicators, will determine whether anger monitoring systems will need to differentiate between sources of anger in their analysis.**

## 1. Introduction

Studies aiming to detect drivers' anger mainly reported an increase in cardiac and respiratory rhythms, skin conductance level and response (Kreibig, 2010) and a decrease in visual exploration (Zhang et al., 2016). Anger while driving can also impair attention by degrading the driver's situational awareness (Jeon et al., 2014).

However, these findings are not universally accepted, and it is not uncommon that no variation is observed following anger induction. Driving anger could be linked to driving situations (e.g. being slowed down by a vehicle) or not (e.g. remember emotional memories). In the literature, different inducing procedures could reproduce these two sources of anger. One question arises: are the two anger induced are linked to the same effects on ocular and physiological responses.

To address this question, we set up an experimental protocol on a driving simulator. We collected cardiac, respiratory, electrodermal activities and eye behaviors after inducing an angry (or neutral) state related and/or unrelated to the driving environment. Emotion induction from the driving environment was made by manipulating events occurring during the driving scenario. Emotion induction unrelated to the driving environment was made thanks to the autobiographical recall technique consisting of asking participants to write and then constantly think about a personal emotional experience. Since anger triggered by driving situations also occurs during autonomous driving (Techer et al., 2019), we focused solely on autonomous mode to rule out any impact from participants' motor actions on the observed differences.

We hypothesized that the indicators relative to the feeling of anger (increased heart rate, respiratory rate, skin conductance and the narrowing of visual exploration) would be: (a) present, whatever the source of induction (by the driving environment, by thoughts of anger, or by a combination of both) ; (b) present all along the drive following an unique induction by angry thoughts ; (c) present only during driving events following an unique induction by the driving environment (d) be more intense and last longer following an induction by a combination of the two techniques.

## 2. Method

### 2.1 *Participants*

37 volunteers (24 men, 13 women) with a mean age of 38, 57 (SD = 13,79) took part in this experiment. Out of the 37 participants, 3 were removed from analysis due to simulator and/or eye-tracking issues.

### *2.2 Material & Measures*

The experiment took place in a homemade fixed-based driving simulator. The driving simulator structure is composed of 2 driving seats (driver and passenger), a Logitech G29 steering wheel and pedals set. The driving environment is created using Unity 3D software.

BIOPAC MP160 was used to collect physiological measures at a sampling rate of 500 Hz.

Fovio, a desktop eye tracker was used to capture ocular metrics at a sampling rate of 62 Hz.

### 2.3 *Experimental Design*

A mixed design was used in this experiment to compare the single and combined effects of anger driving events and anger thoughts. Participants were divided into 4 groups:

- Neutral thoughts and Neutral driving events (NN)
- Neutral thoughts and Anger driving events (NA)
- Anger thoughts and Neutral driving events (AN)
- Anger thoughts and Anger driving events (AA)

### 2.4 *Protocol*

After a 5-minutes rest (Baseline), participants were asked to rate their emotional state from 0 to 100 for different emotions (anger, frustration, joy, sadness, pleasure, fear, disappointment, surprise) and from 0 to 9 for valence, arousal and control over their emotional state.

Then, for 10 minutes, depending on their group, they completed either the anger or neutral (daily routine) autobiographical recall task. Once again, they were asked to evaluate their emotional state

Afterwards, depending on their group, participants followed either the anger or the neutral autonomous driving scenario. For both scenarios, participants were instructed to supervise the environment and to think as much as possible about their memory they just wrote. In the anger scenario, 4 events were added (see Table 1). At the end of the drive, participants in NA and AA groups had to rate their level of perceived anger for the 4 driving events they encountered.

Table 1. Description of the driving events used for anger/neutral induction

| Events name | Anger Events Description | Neutral Events Description | Length (s) |
|---|---|---|---|
| Evt1&2 | A vehicle re-accelerates to avoid being overtaken. Then, once overtaken, it accelerates again to stick to the participant's vehicle before finally overtaking and moving ahead. | The vehicle lets itself be overtaken | 60 |
| Evt3 | A vehicle in the left lane is driving well below the speed limit | The vehicle is in the right-hand lane and the speed limit corresponds to its speed | 53 |
| Evt4 | In a long traffic jam, another driver overtaking by the urgency lane | The vehicle is an ambulance | 30 |

## 3. **Results**

Data analysis is still ongoing. Thus only data from subjective evaluations, and ocular metrics are presented here.

### 3.1 *Anger throughout the drive : subjective results*

The effectiveness of the autobiographical recall technique and driving event to induce anger was checked by using a 4 Groups (AA, AN, NA, NN) * 3 Moments (Baseline, Post Induction, and Post Events) mixed ANOVA design for all likert scales and for the valence, arousal and control scales. Significant results concerning anger scores are illustrated in Figure 1. In order to check our assumptions, pairwise comparisons (with Bonferroni corrections) were made between groups :

- Anger events effects : NA vs. NN groups
- Anger thoughts effects : AN vs. NN groups
- Anger events*thoughts effects : AA vs. NA and AA vs. AN groups

Hypothesis (a) and (b) are verified. Indeed, the ANOVA revealed a main effect of Group ($F(3, 31) = 3.972$, $p = 0.017$, $\eta2 = 0.202$) Moment ($F(2,62) = 16.237$, $p < 0.001$, $\eta2 = 0.152$) and interaction Group and Moment ($F(6,62) = 5.241$, $p < 0.001$, $\eta2 = 0.148$). Pairwise multiple comparisons showed that the score of anger is higher at post induction in AA (M = 42.9; SD = 29.0) and AN (M = 33.5; SD = 25.5) than at baseline in AA : (M = 5.0; SD = 6.4) and AN (M = 6.1; SD = 13.2).

Moreover, the score of anger remained high in Post Events only in the AA group (M = 37.2 ; SD = 21.1). This result, in line with the hypothesis (d), shows that only participants who had a combination of both induction techniques retained a high level of anger at the end of the drive.

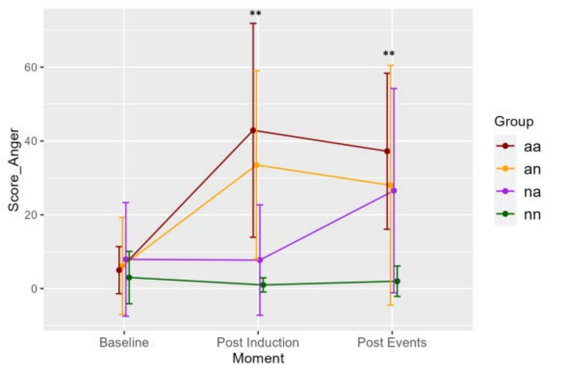



Figure 1. Mean scores of anger declared after the baseline, after induction and post-events in aa, an, na and nn groups. The error bars denote the standard deviation. **$p \leqslant 0.01$

### 3.2 *Anger during driving events: ocular metrics*

To precise the analysis, we specifically checked the evolution of ocular metrics during the driving events.

We measured the ocular effects of both induction techniques by using a 4 groups (NN, NA, AN, AA) * 3 Moments (Evt1 and2, Evt3, Evt4) mixed ANOVA design. It revealed significant differences between groups only during the Evt1&2 time window. However, the results are contrary to our assumption. Participants in the control group (NN) made fewer fixations, and saccades, have more fixation duration time and have a narrower vertical gaze exploration than others. As an illustration (see Figure 2), the ANOVA for the mean fixation duration revealed a main effect of Group ($F(3,28) = 3.200$, $p = 0.038$, $\eta 2 = 0.181$) and an interaction effect of Group and Moment ($F(6,56) = 3.849$, $p = 0.003$, $\eta 2 = 0.128$). Participants from the control group (NN) made longer fixation duration during Evt1&2 (M = 1096.0; SD = 345.0) than NA (M = 505.0; SD = 136.3) and AN groups (M = 701.9 ; SD = 350.4).

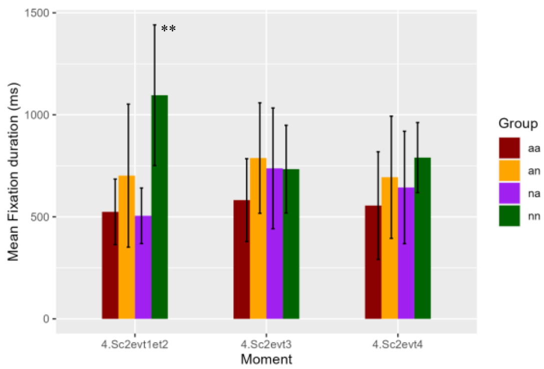



Figure 2. Mean fixation duration for events 1and2, 3 and 4. The error bars denote the standard deviation. $^{**}p \leqslant 0.01$

## 4. Discussion

In this experiment, we attempted to evaluate if a different source of the driver's anger (whether related to the driving environment or not) is associated with different ocular and physiological behaviours. For this purpose, First, 37 participants were asked to think constantly about a neutral or anger-related personal event. Then they drove in an autonomous mode and either encountered anger or neutral driving events.

First, based on subjective data, both techniques succeeded to induce anger and the combination of both procedures worked better to induce an effect that lasts.

Moreover, participants with the neutral autobiographical recall task and neutral driving events showed a narrower vertical gaze exploration, fewer fixations and saccades and an increase in fixation durations than participants induced in anger. Unlike the study of Zhang et al., (2016) we did not find that anger-induced participants had a narrower horizontal visual exploration.

These differences can be interpreted by considering the nature of our neutral autobiographical recall task. Participants were asked to write and think about their daily lives. Thinking about unrelated thoughts while driving has been linked with an increase in gaze fixation (Pepin et al., 2018).

The solely use of eye metrics was not sufficient in order to discriminate the source of the driver's anger. The inclusion of physiological data should enable us to refine the relevant indicators.

## 5. Conclusions

In a simulator study, anger was induced by driving events and/or by thoughts unrelated to the driving environment. Analysis from subjective data highlighted that both induction techniques worked well separately and better when combined. Analysis from the ocular metric revealed that participants induced in neutral have more fixations, more fixation duration and a narrowed visual exploration. However, no difference was found between both induction techniques. The analysis of other physiological indices will enable us to go further and answer the question of whether it is necessary to distinguish between the two types of induction for the development of algorithms to detect anger at the wheel.

## 6. Acknowledgments


The analysis of this research was conducted with French financial support of BPI in the SERENITE project.


## References


Jeon, M., Walker, B. N., & Gable, T. M. (2014). Anger effects on driver situation awareness and driving performance. *Presence: Teleoperators and Virtual Environments*, 23(1), 71-89. https://doi.org/10.1162/PRES_a_00169

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological psychology*, 84(3), 394-421. https://doi.org/10.1016/j.biopsycho.2010.03.010

Pepin, G., Malin, S., Jallais, C., Moreau, F., Fort, A., Navarro, J., ... & Gabaude, C. (2018). Do distinct mind wandering differently disrupt drivers? Interpretation of physiological and behavioral pattern with a data triangulation method. *Consciousness and cognition*, 62, 69-81. https://doi.org/10.1016/j.concog.2018.04.006

Techer, F., Ojeda, L., Barat, D., Marteau, J. Y., Rampillon, F., Feron, S., & Dogan, E. (2019). Anger and highly automated driving in urban areas: The role of time pressure. *Transportation research part F: traffic psychology and behaviour,* 64, 353-360. https://doi.org/10.1016/j.trf.2019.05.016

Zhang, T., Chan, A. H. S., Ba, Y., & Zhang, W. (2016). Situational driving anger, driving performance and allocation of visual attention. *Transportation Research Part F: Traffic Psychology and Behaviour, 42*, 376-388. https://doi.org/10.1016/j.trf.2015.05.008

# Inducing and Regulating Anger and Sadness in Delegated Driving: An In-Car Experiment

Jordan Maillant [12*], Florence Mazères[2], Yannick Lavril[2], Alexandra Fort[2], Hélène Tattegrain [2], Stéphanie Dabic [1], Christophe Jallais [2]

[1] Valeo Switches & Smart Controls Product Line, Rue Jules VERNE - Vetraz Monthoux
BP 509 - 74106 Annemasse Cedex, FRANCE
[2] Univ Gustave Eiffel, Univ Lyon, LESCOT, F-69675 Lyon, France
[*]jordan.maillant@valeo.com

**Abstract:**

**In delegated driving, the driver adopts a position of supervisor. While automation can influence attentional and emotional states, the driver must remain attentive and serene to be able to regain control if necessary. This study, carried out in a real car under delegated driving conditions, aimed to detect when the mind wanders towards negative (anger or sadness) thoughts and to help the driver regulate these negative states with the help of cardiac coherence. 46 volunteers took part in the experiment. They were divided into 2 groups: with and without regulation. Emotions (anger, sadness and serenity) were within-subject factors. Subjective data highlighted the effectiveness of the mood induction procedure (autobiographical recall) in inducing negative states. Subjective data highlighted the effectiveness of the mood induction procedure (autobiographical recall) in inducing negative states, but did not allow us to assess the effectiveness of the regulation technique implemented (cardiac coherence). Further analyses of physiological data should enlighten this potential effect.**

## 1. Introduction

Existing literature suggests that drivers who have experienced a recent stressful emotional event are more likely to have at-fault accidents (Lagarde et al., 2004). Similarly, drivers who tend to focus on negative thoughts report being more likely to engage in aggressive behavior (Suhr, 2016; Love et al., 2022; Stephens et al., 2023). In delegated driving, the driver is expected to remain attentive to the environment and ready to regain control if necessary or requested. However relying on the system can lead to under-activation, which can hinder the mobilization of attentional resources (Young & Stanton, 2002) and leave room for mind wandering (Gouraud et al., 2018). Given the positive correlation between the emergence of mind wandering and the emergence of negative thoughts while driving (Walker & Trick, 2018), delegated driving would be a fertile ground for inattention linked to negative thoughts. It is therefore important to monitor and regulate the driver's emotional state during the delegated driving. The aim is to detect when the mind is wandering towards negative thoughts and help the driver return to a serene state.

To achieve this aim, we set up an experiment that was carried out in a closed-circuit with a simulated autonomous car. Negative (anger and sadness) and neutral (serenity) emotions were induced. Breathing exercise was used to help them defocus their emotional thoughts and refocus on themselves and their environment. Cardiac, respiratory and electrodermal activities were recorded all along the experiment. To the best of our knowledge, this study is the first to address these objectives in delegated driving on a real vehicle, using a neuroergonomic approach.

## 2. Method

### *2.1. Participants*

46 volunteers (26 men and 20 women) with a mean age of 41.0 (SD = 11.2) took part in this experiment. Of the 46 participants, 2 were excluded due to technical problems and 1 due to medication that can act on emotional regulation. Thus, the data analysis included 43 participants. All participants gave their consent and were rewarded with 120 euros.

### *2.2. Road track description*

The experiment took place on a 1.755 km ring track at the Transpolis test center for road safety located at Amberieu en Bugey in France.

### *2.3. Material & Measures*

Participants were installed at the driver's seat of a Mercedes Class E AMG Break. The car was modified in order to provide a Wizard Of Oz (WOO) autonomous driving.

BIOPAC MP160 was used to collect physiological measures (ECG, RSP, EDA) at a sampling rate of 500 Hz.

### *2.4. Experimental Design*

A mixed approach design was conducted. The emotion induced (anger, sadness, serenity) was an intra-subject factor and the presence or absence of emotion regulation was a between-subject factor.

### *2.5. Protocol*

The experiment lasted about 4 hours. It consisted of a series of 3 experimental sessions in the car with a 20-minute break between sessions.

Each experimental session lasted about 30 minutes. It started with a 5-min rest (*Baseline*) and then involved inducing and regulating/not regulating an emotion (anger, sadness or serenity). Emotions were *induced* by the autobiographical recall technique consisting of asking participants to write in detail an anger, sad or serene personal event. The emotion induced was maintained during the experiment by visual instructions displayed on a tablet (*Recall Induction*) asking them to think again about this. The *regulation* (breathing exercise) consisted of a rhythm composed of 5 seconds (inhalation) with vibrations and 10 seconds without vibrations (exhalation).

Five times during the experiment (repeated measure), participants were asked to rate their emotional state from 0 to 100 according to 4 emotions: sadness, joy, serenity, anger as well as on the dimension of control (i.e. the level they exert on their emotional state).

In line with the literature (Kreibig, 2010), we assumed that anger induction from thoughts would increase cardiac, respiratory and electrodermal activities, while sadness would decrease them and serenity would not. The declared scores of anger and sadness would be higher right after the induction and be maintained after the phase of recall induction.

## 3. Results

Data analysis is still in progress. Consequently, only data relating to subjective evaluations following emotional inductions are presented here.

Delta scores were calculated by subtracting the scores for each emotion from those reported in the baseline.

To evaluate the effectiveness of the emotion induction procedure, we conducted, for all scales, a mixed ANOVA with 3 (Emotion_Induced (within factor): Anger, Sadness, Serenity) * 4 (Moment (within factor): Induction, Post-induction, Post-Regulation, Post-Drive) * 2 (Group (between factor): Regulation, No Regulation).

In line with our assumptions, for each emotion scale, pairwise comparisons (with Bonferroni corrections) were made between emotions induced for each moment.

Significant findings concerning anger and sadness scores are presented here and illustrated in Figure 1.

### *3.1. Anger Scores*

The ANOVA with Greenhouse-Geiser corrections revealed a main effect of Moment ($F(2.20,90.34) = 17.952$, $p < 0.001$, $\eta 2 = 0.022$) Emotion_Induced ($F(1.61, 66.21) = 9.879$, $p < 0.001$, $\eta 2 = 0.120$) and an interaction effect Moment: Emotion_Induced ($F(3.75, 153.82) = 5.678$, $p < 0.001$, $\eta 2 = 0.012$).

In the non-regulated group, for anger-induced participants, delta scores of anger are higher for every moment : induction (M = 27.5; SD = 29.8), post induction (M = 26.9; SD = 34.9), post regulation (M= 21.1; SD = 31.8) and post driving (M = 13.7 ; SD = 26.1) than serenity-induced participants (induction : M = -0.9 ; SD = 4.9) (post induction : M = -1.1 ; SD = 4.7) (post regulation : M = -1.1 ; SD = 5.4) (post driving : M = -0.9 ; SD = 4.9). Moreover, the scores of anger after post induction and post regulation are higher for anger-induced participants than sadness-induced participants (post induction: M = 5.4; SD = 16.4) (post regulation: M = 5.5 ; SD = 13.4).

In the regulated-group, only a difference between anger and serenity inductions is observed after induction (anger: M = 25.4 ; SD = 33.8) (serenity: M = 1.3 ; SD = 24.0).

### *3.2. Sadness Scores*

The ANOVA with Greenhouse-Geiser corrections revealed a main effect of Moment (F(1.94,79.37) = 39.942, $p < 0.001$, $\eta2 = 0.066$) Emotion_Induced (F(2.00, 82.00) = 10.999, $p < 0.001$, $\eta2 = 0.121$) and an interaction effect Moment: Emotion_Induced (F(3.12,127.92) =17.362, $p < 0.001$, $\eta2 = 0.053$).

In the non-regulated group, for sadness-induced participants, delta scores of sadness are higher after induction (M = 40.5 ; SD = 25.5), post induction (M = 22.4 ; SD = 22.7) and post regulation (M = 14.2 ; SD = 17.4) than following a serenity induction (induction: M = -0.5 ; SD = 3.7) (post induction: M = -0.3 ; SD = 5.2) (post regulation: M = -1.5 ; SD = 4.1). Moreover, delta scores of sadness are also higher after induction compared to anger-induced participants (M = 18.3 ; SD = 22.7). Additionally, anger induction arised sadness scores. The delta scores of sadness are higher following anger induction in every moment: induction (M = 18.3 ; SD = 22.7), post induction (M = 17.7; SD = 24.5), post regulation (M = 14.4 ; SD = 25.4), post driving (M = 11.5 ; SD = 22.1) than following the induction of serenity.

In the regulated group, the delta scores of sadness are only higher after sadness induction (M = 40.8 ; SD = 31.8) than for anger-induced (M = 12.9 ; SD = 27.8) and serenity-induced participants (M = 11.8 ; SD = 25.9).

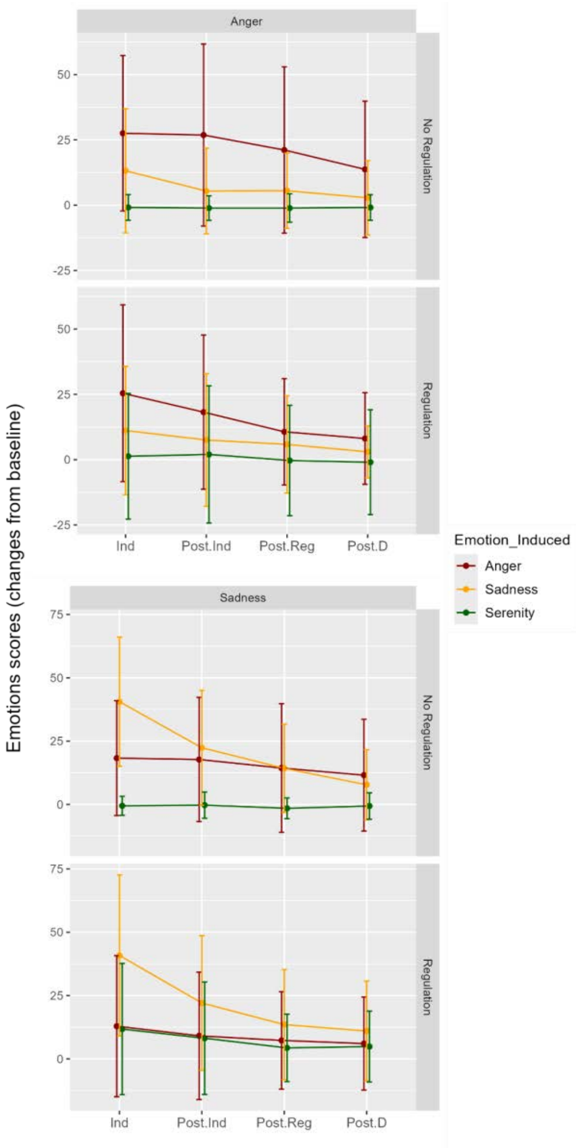



Figure 1. Delta scores of anger and sadness following anger, sadness or serenity inductions declared in both groups (no regulation and regulation) after : induction (Ind), Post induction (Post.Ind), regulation (Post.Reg) and at the end (Post Driving). The error bars denote the standard deviation.

## 4. Discussion

Significant and lasting changes in sadness and anger scores were measured after induction but differences were observed across groups. For instance, in the non-regulated group, participants induced with anger consistently exhibited higher scores of anger compared to serenity-induced participants across all moments. Conversely, in the regulated group, while the induction of sadness or anger indeed induced an increase in the respective sadness and anger scores, this effect did not last over time. In addition, recalling participants' negative thoughts elicited a variety of emotions. Notably, sadness induction also contributes to a significant increase in anger in the unregulated group. The objectives here concerned not only the induction procedure's effectiveness but also the possibility of coping with emotion using cardiac coherence. The results did not allow us to conclude about the efficiency of the technique. Indeed, it seemed that the use of automation decreased the effectiveness of mood induction. These findings underscore the nuanced interplay between emotional induction procedures and regulatory mechanisms, shedding light on the complex dynamics of emotional experiences in delegated driving contexts. In our study, participants had to drive under automation mode for 17 minutes in a poorly stimulating environment (i.e., tracks). These driving conditions could lead to reduced vigilance or mind-wandering, potentially affecting their responsiveness to emotional induction procedures.

## 5. Conclusions

From this study, carried out on a real car, initial analyses from subjective data revealed the possibility of inducing a negative emotional state under delegated driving conditions. The subsequent analysis of physiological data should provide a better understanding of the complex interplay between emotional reactions and physiological states in the context of delegated driving. This holistic approach should provide valuable insights into how to improve safety and comfort by controlling emotions and attention during delegated driving.

## 6. Acknowledgments

The analysis of this research was conducted with French financial support of BPI in the SERENITE project.

## References

Gouraud, J., Delorme, A., & Berberian, B. (2018). Influence of automation on mind wandering frequency in sustained attention. *Consciousness and cognition*, 66, 54-64.

Lagarde, E., Chastang, J. F., Gueguen, A., Coeuret-Pellicer, M., Chiron, M., & Lafont, S. (2004). Emotional stress and traffic accidents: the impact of separation and divorce. *Epidemiology*, 15(6), 762-766.

Love, S., Kannis-Dymand, L., Davey, J., & Freeman, J. (2022). Metacognition, rumination and road rage: an examination of driver anger progression and expression in Australia. *Transportation research part F: traffic psychology and behaviour*, *84*, 21-32.

Stephens, A. N., Collard, J., & Koppel, S. (2023). Don't sweat the small stuff; anger rumination and lack of forgiveness are related to aggressive driving behaviours. *Current Psychology*, 1-12.

Suhr, K. A. (2016). Mulling over anger: Indirect and conditional indirect effects of thought content and trait rumination on aggressive driving. *Transportation research part F: traffic psychology and behaviour*, 42, 276-285.

Young, M. S., & Stanton, N. A. (2002). Attention and automation: new perspectives on mental underload and performance. Theoretical issues in ergonomics science, 3(2), 178-194.

Walker, H. E., & Trick, L. M. (2018). Mind-wandering while driving: The impact of fatigue, task length, and sustained attention abilities. *Transportation research part F: traffic psychology and behaviour*, 59, 81-97.

# DDI2024 MICHIGAN

THE 9TH INTERNATIONAL CONFERENCE ON
DRIVER DISTRACTION AND INATTENTION