frontiers | Research Topics

Cognitive mechanisms for safe road traffic systems

Edited by
Giovanni Vecchiato, Lewis L. Chuang and Christer Ahlström

Published in
Frontiers in Neuroergonomics

ISSN 1664-8714
ISBN 978-2-88976-151-7
DOI 10.3389/978-2-88976-151-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Cognitive mechanisms for safe road traffic systems

Topic editors
Giovanni Vecchiato — Institute of Neuroscience, National Research Council of Italy, Italy
Lewis L. Chuang — Ludwig Maximilian University of Munich, Germany
Christer Ahlström — Swedish National Road and Transport Research Institute, Sweden

Citation
Vecchiato, G., Chuang, L. L., Ahlström, C., eds. (2022). *Cognitive mechanisms for safe road traffic systems*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-151-7

Table of contents

frontiers | Frontiers in Neuroergonomics

EDITORIAL
published: 25 April 2022
doi: 10.3389/fnrgo.2022.897659

Editorial: Cognitive Mechanisms for Safe Road Traffic Systems

Giovanni Vecchiato [1†], Christer Ahlström [2,3†] and Lewis L. Chuang [4†]*

[1] Institute of Neuroscience, National Research Council of Italy, Parma, Italy, [2] Swedish National Road and Transport Research Institute (VTI), Linköping, Sweden, [3] Department of Biomedical Engineering, Linköping University, Linköping, Sweden, [4] Department Humans and Technology, Institute for Media Research, Faculty of Humanities, Chemnitz University of Technology, Chemnitz, Germany

Keywords: automated driving, driving behavior, road safety, brain signals, eye tracking, EMG, autonomic signals

OPEN ACCESS

Edited by:
Klaus Gramann,
Technical University of Berlin, Germany

Reviewed by:
Thea Radüntz,
Federal Institute for Occupational Safety and Health, Germany
Mathias Vukelić,
Fraunhofer Institute for Industrial Engineering, Germany

***Correspondence:**
Giovanni Vecchiato
giovanni.vecchiato@in.cnr.it

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: *16 March 2022*
Accepted: *04 April 2022*
Published: *25 April 2022*

Citation:
Vecchiato G, Ahlström C and Chuang LL (2022) Editorial: Cognitive Mechanisms for Safe Road Traffic Systems.
Front. Neuroergon. 3:897659.
doi: 10.3389/fnrgo.2022.897659

The Editorial on the Research Topic

Cognitive Mechanisms for Safe Road Traffic Systems

Human behavior is often cited as the primary contributing factor to road accidents—over 90% of all crashes are attributed to "human error" (Singh, 2015). This implicitly suggests that accidents could be avoided if only drivers behaved better and has, thus, fuelled enthusiasm for (semi-)automated vehicles, which do not suffer from human frailty and are more likely to follow the rules. Nonetheless, this perspective is flawed. First, most road users strive to avoid road accidents. Second, fatalities persist even with (semi-)automated vehicles and it remains unclear if increased adoption of more automation will change the situation at all (Mueller et al., 2021). Modern perspectives suggest that "human error" is a product of not only individual behavior but the system that we operate within (Read et al., 2021). An individual cannot be understood without taking into account their relationship with the working environment. Safe vehicles are those that enable drivers to act with a minimal margin for unintended error while ensuring that road traffic systems cater to user autonomy. Even if automation can mitigate driving-related risks, it will simply present new challenges that can only be anticipated by first understanding the cognitive mechanisms associated with operating in road traffic systems. To this aim, it is vital that we possess better tools to understand, measure and monitor human behavior and the corresponding cerebral activity across diverse road scenarios, including those that do not generate overt behavior.

This Research Topic invited manuscripts that covered modeling, behavioral and neurophysiological measures investigating conventional and (semi-)automated driving with the goal to develop a safe human-centric road traffic system. The submitted contributions responded to this challenge in various ways, including state-of-the-art physiological measures to assess the driver's mental state, theoretical and empirical methodological approaches to advance the present knowledge, and challenges in vehicle automation.

Several papers in this article Research Topic emphasize that fatigue, cognitive load, inattention, and stress are multidimensional constructs that must be interpreted within a context, taking both endogenous and exogenous factors into account. This is important not only for accurate estimation of the driver's state, but also when selecting appropriate countermeasures.

Here, the review by Chong and Baldwin addresses the underlying mechanisms of fatigue and how different types of fatigue arise. For example, active fatigue is related to long-term neuronal potentiation and local sleep, whereas passive fatigue disturbs the interplay between the dorsal attention network and the default mode network. Circadian effects can further moderate both passive and active fatigue, linking the two to a third category of fatigue, namely sleep-related fatigue. Discriminating for different types of fatigue is a crucial first step for adopting appropriate countermeasures: active fatigue should be countered by reduced task demand, passive fatigue by increased task demand, and sleep-related fatigue by sleep and recuperation.

The effects of task demand on cognitive load is addressed by Nilsson et al. Their approach accounts for the many complexities that arise when estimating cognitive load in a realistic situation, such as traffic. First, they characterize cognitive load as a multidimensional construct that consists of many mental responses to added task demand. Second, they highlight that cognitive load does not occur in isolation but is part of a complex response to task demands in a specific context. Finally, physiological measures typically correlate with more than one mental state, thus limiting the inferences that can be made from any individual state.

Kerautret et al. provide a systematic review that guides the selection of appropriate physiological measures for quantifying acute stress. Once again, physiological response to a driver state change is characterized as a multidimensional construct, where several measures, including heart rate, R-R intervals and pupil diameter, respond to driver stress levels. When aiming to improve the practical usefulness of stress detection devices, it is important to start considering the context where the stress level is measured. Increased sympathetic activity is a reasonable response to a critical event or complex unfamiliar environments, and it is important to realize that a temporarily elevated stress level is not necessarily a bad thing.

What is appropriate and inappropriate behavior? More importantly, Ahlström et al. demonstrate that this deceptively simple classification depends on context in this article that describes eye movements and how they relate to attentive driving. Essentially, it is not just where drivers look, but also why and what else they can see and where they do not look. The authors suggest a glance analysis approach that classifies glances based on their purpose, thus accounting for context or motivation behind eye movements.

The present article Research Topic also comprises a series of theoretical and empirical works that leverage methodological aspects to foster the knowledge on fundamental brain processes for driving research.

Kujala and Lappi outline a predictive processing approach for studying attentional demand and inattention in driving, based on neuro-inspired theories of uncertainty processing and experimental research that combine brain imaging, visual occlusion and computational modeling. This approach improves the definition and detection of inattentive driving as a step toward designing attention monitoring systems for conventional and semi-automated driving.

Vecchiato provides a perspective on the technological maturity of hybrid systems, which could improve the predictive power of a single neurophysiological measurement. Electroencephalography (EEG) is often constrained by robustness, comfortability, and high data variability affecting the decoding performance in driving scenarios. Hence, additional peripheral signals can be combined with EEG for increasing replicability and the overall performance of the brain-based action decoder.

Getzmann et al. explored the ecological and internal validity of round-the-ear electrodes (cEEGrids) measurements. Their longitudinal study returns consistent modulations in the alpha and theta bands, along with driving speed and steering wheel angular velocity reflecting the complexity of the driving task between the two measurements. Overall, the reliability and ecological validity of cEEGrid electrodes were satisfactory in the context of driving-related parameters.

Among the many issues related to the safe use of driving automation, attention drift due to the modulation of internal sources could play an important role in the emergence of out-of-the-loop (OOTL) situations and associated performance problems (Merat et al., 2019).

Gouraud et al. address the possibility of the gradual emergence of attentional decoupling and the differences created by the sensory modality used to convey targets using EEG measurements. Their results underline the complex influence of perceptual decoupling on operators' behavior and EEG measures.

Sensory skills can be augmented through training and technological support. Exploiting this, Sakai et al. used fMRI to compare brain responses to auditory cues for self-localization, modulated by a sensory augmentation training in a simulated driving environment. Their results suggest that the use of auditory cues for self-localization during locomotion relies on multimodality in higher-order somatosensory, rather than visual, areas.

The complexity of autonomous navigation increases due to the lack of road signs and pedestrians' presence. The work by Petit et al. deals with the perception of collision risk from the viewpoint of a passenger sitting in the driver's seat. Such users delegate total control of their vehicle to an autonomous system and this article investigated the subjective risk assessment with a system based on the measurement of the electrodermal activity. The results demonstrate that reducing safety margins increases risk perception.

Considerable evidence suggest that humans may interact differently with autonomous vehicles (AVs) as compared to human-driven vehicles (HVs). Unni et al. investigated whether participants would value interactions with AVs differently compared to HVs, and if these differences can be characterized in terms of behavior and brain responses. Using hemodynamic response features from whole-head fNIRS, they could predict whether participants decided to turn in front of HVs or AVs in the decision-making phase. The insights provided here may be useful for developing driver assistance systems to assess interactions in future mixed traffic environments involving AVs and HVs.

Finally, Fredriksson et al. provide a roadmap for the development of future Occupant Status Monitoring (OSM) in the EuroNCAP protocol. This considers a range of known and emerging safety risks, including driving while intoxicated by alcohol or drugs, cognitive distraction, and the driver engagement requirements for supervision and take-over performance with assisted and automated driving features.

The works collected by this Research Topic describe a wide range of challenges that have to be addressed as we improve on our knowledge of cognitive mechanisms, relevant for the design of safe road traffic systems. In view of the strong interest in the academic and industrial fields, we believe that the present Research Topic will increase rigor and reproducibility in driving research.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

LC is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 416228727 - SFB 1410. GV is supported by a reseach agreement between CAMLIN Limited and IN-CNR.

REFERENCES

Merat, N., Seppelt, B., Louw, T., Engström, J., Lee, J. D., Johansson, E., et al. (2019). The "Out-of-the-Loop" concept in automated driving: proposed definition, measures and implications. *Cogn. Technol. Work* 21, 87–98. doi: 10.1007/s10111-018-0525-8

Mueller, A. S., Reagan, I. J., and Cicchino, J. B. (2021). Addressing driver disengagement and proper system use: human factors recommendations for level 2 driving automation design. *J. Cogn. Eng. Decision Making* 15, 3–27. doi: 10.1177/1555343420983126

Read, G. J., Shorrock, S., Walker, G. H., and Salmon, P. M. (2021). State of science: evolving perspectives on 'human error'. *Ergonomics* 64, 1091–1114. doi: 10.1080/00140139.2021.1953615

Singh, S. (2015). *Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey* (No. DOT HS 812 115).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroergonomics

ORIGINAL RESEARCH
published: 01 April 2021
doi: 10.3389/fnrgo.2021.625343

Mind Wandering Influences EEG Signal in Complex Multimodal Environments

Jonas Gouraud[1]*, *Arnaud Delorme*[2] *and Bruno Berberian*[1]

[1] Systems Control and Flight Dynamics Department, Office National d'Etudes et de Recherche Aérospatiales, Salon de Provence, France, [2] Center of Research on Brain and Cognition (UMR 5549), Centre National de Recherche Scientifique, Toulouse, France

The phenomenon of mind wandering (MW), as a family of experiences related to internally directed cognition, heavily influences vigilance evolution. In particular, humans in teleoperations monitoring partially automated fleet before assuming manual control whenever necessary may see their attention drift due to internal sources; as such, it could play an important role in the emergence of out-of-the-loop (OOTL) situations and associated performance problems. To follow, quantify, and mitigate this phenomenon, electroencephalogram (EEG) systems already demonstrated robust results. As MW creates an attentional decoupling, both ERPs and brain oscillations are impacted. However, the factors influencing these markers in complex environments are still not fully understood. In this paper, we specifically addressed the possibility of gradual emergence of attentional decoupling and the differences created by the sensory modality used to convey targets. Eighteen participants were asked to (1) supervise an automated drone performing an obstacle avoidance task (visual task) and (2) respond to infrequent beeps as fast as possible (auditory task). We measured event-related potentials and alpha waves through EEG. We also added a 40-Hz amplitude modulated brown noise to evoke steady-state auditory response (ASSR). Reported MW episodes were categorized between task-related and task-unrelated episodes. We found that N1 ERP component elicited by beeps had lower amplitude during task-unrelated MW, whereas P3 component had higher amplitude during task-related MW, compared with other attentional states. Focusing on parieto-occipital regions, alpha-wave activity was higher during task-unrelated MW compared with others. These results support the decoupling hypothesis for task-unrelated MW but not task-related MW, highlighting possible variations in the "depth" of decoupling depending on MW episodes. Finally, we found no influence of attentional states on ASSR amplitude. We discuss possible reasons explaining why. Results underline both the ability of EEG to track and study MW in laboratory tasks mimicking ecological environments, as well as the complex influence of perceptual decoupling on operators' behavior and, in particular, EEG measures.

Keywords: out of the loop, mind wandering, automation, vigilance, attentional decoupling, sensory modalities

Edited by:
Lewis L. Chuang, Ludwig Maximilian University of Munich, Germany

Reviewed by:
Sarah Blum, University of Oldenburg, Germany
Shaibal Barua, Mälardalen University College, Sweden

***Correspondence:**
Jonas Gouraud
contact@jonasgouraud.com

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: *02 November 2020*
Accepted: *23 February 2021*
Published: *01 April 2021*

Citation:
Gouraud J, Delorme A and Berberian B (2021) Mind Wandering Influences EEG Signal in Complex Multimodal Environments. Front. Neuroergon. 2:625343. doi: 10.3389/fnrgo.2021.625343

INTRODUCTION

Context

The last decade has seen important research toward road automation (Badue et al., 2019). Promised as a revolution for users to gain flexibility, leisure time, and safety (Harb et al., 2018; Correia et al., 2019), self-driving cars have nonetheless several important technological gaps that must be filled before becoming a reality. On the way toward level 5 automation (full automation anywhere, see SAE International, 2018), teleoperation could represent an important trade-off to maintain safety while developing system capabilities. Teleoperation, literally operating a vehicle at a distance, is already used in environments unreachable or dangerous to humans, such as war theaters, nuclear environments, and space (Lichiardopol, 2007). Tomorrow, teleoperation could be performed by algorithms in the cloud and allow any vehicle to reach level 5 automation with minimal modifications (Zhang, 2020). However, the technology could also use human intervention today to enhance partial automation and widen its operational design domain (Kang et al., 2018). Operators would then monitor a set of vehicles, taking control whenever necessary, such as in the event of snow or in an emergency. Specifically, an important advantage of human teleoperation is the assumption that there could be more vehicles to monitor than operators, as not all vehicles would require assistance at the same time (Zhang, 2020).

Aside from technical challenges like latency (Neumeier et al., 2019), the possibility of jumping into a specific situation only when needed raises important interrogations regarding the ability of operators to assume manual control when needed. Humans would then only have to monitor, presumably ever-alert, for deviations and problems. Situations where operators are supervising automated control loop are called out-of-the-loop (OOTL) situations (Norman and Orlady, 1988; Endsley and Kiris, 1995). Unfortunately, OOTL situations reduce the operators' ability to intervene, if necessary, and to assume manual control, i.e., to come back in the control loop (Kurihashi et al., 2015; Louw et al., 2015a; Naujoks et al., 2016). Supervisors at this point seem dramatically powerless to diagnose the situation, determine the appropriate solution, and execute it before the accident happens. Accident reports may contain the terms "total confusion" (National Transportation Safety Board, 1975, 17; Bureau d'Enquête et d'Analyse, 2002, 167), "surprise effect" (Bureau d'Enquête et d'Analyse, 2012a, 44, 2016, 10), or "no awareness of the current mode of the system" (Bureau d'Enquête et d'Analyse, 2012b, 178). These negative side effects on overall performance are commonly referred to as OOTL performance problems.

Nowadays, it is assumed that OOTL performance problem is fundamentally a matter of human–automation interaction arising from both operators' internal states and system properties, which ultimately spoils performance (Berberian et al., 2017). From this definition, one way to mitigate related performance drops may be to monitor operators' internal states and look for precursors to OOTL performance problems. Among others, it has been demonstrated that non-challenging tasks, such as passive monitoring of automation, can promote episodes of mind wandering, whereby attention drifts away from the task at hand (Smallwood et al., 2008; Durantin et al., 2015; Smallwood and Schooler, 2015; Gouraud et al., 2018a,b; Dehais et al., 2020). Mind wandering (MW) is a family of experiences unrelated to the here and now (Seli et al., 2018). When MW happens during a task, it moves operators' minds away from their tasks toward matters not directly related to their current works. Although such uncontrolled thoughts could be beneficial as long-term planning and mind refreshment (McMillan et al., 2013; Ottaviani and Couyoumdjian, 2013; Terhune et al., 2017), it may thwart short-term performances (He et al., 2011; Galera et al., 2012; Cowley, 2013; Casner and Schooler, 2014; Dündar, 2015). Therefore, real-time tracking of MW is an important goal within safety-critical industries, particularly when automation supervision fills a significant part of the job. Indeed, real-time tracking of internal states like MW would allow detecting problems before performance drops and accidents happen. However, a better understanding of the emergence of this attentional decoupling remains essential to achieve such a goal. This is precisely the objective of this study.

Emergence of Attentional Decoupling

Many physiological tools have already demonstrated sensitivity to several aspects of MW; however, electroencephalography (EEG) is among the most promising. EEG signal has already helped uncover an important facet of MW: attentional decoupling. People subject to MW experience a drop in the cortical processing of the external environment, as their attention is redirected to inner thoughts (Schooler et al., 2011). Neurologically, attentional decoupling is characterized by weaker neuronal responses to external stimuli and greater deactivation of the regions dedicated to their processing. During GO/NOGO tasks, researchers (Kam et al., 2011) showed that the amplitude of P1, N1, and P3 components (respectively associated with visual perception, auditory perception, and external stimuli processing) were all lower during task-unrelated MW. This effect held true whether stimuli were the SART (Sustained Attention to Response Task) stimuli or irrelevant to the task. Such results were replicated in two other settings: a time-estimation task (Kam et al., 2012) and during monotonous manual driving (Baldwin et al., 2017). It was also highlighted through ERPs that attentional decoupling involved lower emotional reactions (Kam et al., 2014). Experiments also uncovered the signature of MW on alpha waves in occipital, i.e., visual stimuli processing, areas (O'Connell et al., 2009; Braboszcz and Delorme, 2011; Baird et al., 2014; Atchley et al., 2017; Baldwin et al., 2017; Arnau et al., 2020), although the exact way is still debated as explicated in the next sections. Nevertheless, changes in alpha activity during MW are in line with the alpha band being involved in the deactivation of the concerned areas (Bonnefond and Jensen, 2012; Benedek et al., 2014; Villena-González et al., 2016).

Factors Influencing Attentional Decoupling

Even though MW has a strong influence on the neuronal signal, the factors modulating the attentional decoupling remain unidentified. A first important question is the exact degree of attentional decoupling. Put differently, do all MW have the same potential for attentional decoupling? Is "depth" a feature of MW episodes? Several studies provide insight into depth as a feature of MW episodes. Cheyne et al. (2009) used a SART to investigate the validity of their bi-directional model of inattention. They obtained converging measures supporting three postulated states of inattention: level 1 characterized by more erratic reaction time, level 2 by anticipations, and level 3 by omissions. Following the same path, Schad et al. (2012) detailed the "levels of inattention hypothesis" based on the assumption that our mind processes information sequentially, involving greater complexity at each step. MW could then thwart information processing at different stages, depending on the depth of the episode. While some MW episodes could be superficial, only impacting higher cognition, others could completely decouple from the task by blocking external information encoding and "cascade through the cognitive system" to impact more complex processing (Smallwood, 2011).

A second issue refers to the impact of MW on non-relevant stimulation. It was initially assumed that MW involves a specific impairment in the processing of task-relevant events (e.g., Smallwood et al., 2003, 2004). Studies using ERPs have shown that MW dampens the processing of sensory information, regardless of the relevance of this information to the task (Barron et al., 2011; Kam et al., 2011). However, the fact that MW can impact mechanisms of selective attention does not mean that all stages of sensory processing are turned off. Rather, it signifies that the highlighting of specific sensory inputs for higher levels of cognitive analysis is attenuated. After all, we are able to perform most of our daily tasks without any errors, even during MW episodes. In this context, steady-state responses (SSR) may highlight the exact impact of MW on cognition. An SSR is an evoked potential emerging from external periodical stimulus and whose phase and amplitude remain constant (Picton et al., 2003). Multiple studies have highlighted that in environments with multiple SSR competing for attention, focusing on one SSR increases its amplitude to the detriment of the others (Skosnik et al., 2007; Müller et al., 2009; Saupe et al., 2009a; Diesch et al., 2012; Mahajan et al., 2014). However, it has been shown that this effect is highly dependent on experiments' features: paying attention to a 20-Hz ASSR presented on one ear showed increase amplitude ipsilaterally, but not for a 40-Hz stimulus (Müller et al., 2009); in another study, the attention-competition effect decreased SSR amplitude only when concurrent SSR were presented on the same sensory modality (Porcu et al., 2014). These results highlight the complexity of the different stages of perception and attention, and SSR may help to understand the influence of MW on them. Moreover, if SSR were to be impacted by MW, it could reveal extraordinarily useful to study the features of attentional decoupling. Indeed, it would allow continuous monitoring, contrary to ERPs, while being fully controlled in frequency, in contrast to natural brain waves. O'Connell et al. (2009) has already investigated the influence of lapses of attention on a visual SSR without finding significant results regarding its amplitude. However, they did not use a questionnaire to track MW episodes. To our knowledge, no research has specifically addressed the impact of internally directed attention on SSR amplitude.

Our purpose in this experiment is to evaluate the viability of MW neuronal markers in complex laboratory tasks mimicking automated ecological environments, as well as help to characterize features of the attentional decoupling in these environments. Our hypotheses are (1) the evolution of MW can be tracked in complex environments through a decrease in ERPs and ASSR amplitude coupled with an increase in alpha power during MW episodes compared with focus moments, (2) MW attentional decoupling demonstrates a gradual nature on EEG measures (ERP, alpha, ASSR) correlated to the proximity of the thoughts content to the task at hand; more precisely, a MW episode with thoughts closer to the immediate environment will have less influence on EEG measures than another MW episodes with thoughts totally unrelated to the here and now.

MATERIALS AND METHODS

Participants

We performed an a priori analysis to estimate the required sample size. Most publications investigating the links between MW and EEG did not report effect size explicitly. However, as repeated-measure ANOVA was often used, we could calculate from these publications Cohen's f using F-value, CI, and degrees of freedom. The lower value computed, which we retained to adopt a conservative view, was 0.54 (Kelley, 2007a,b, 2020; Uanhoro, 2017). We then used G*Power (Faul et al., 2007, 2009) to calculate the sample size, which yielded a minimum of 14 participants.

Eighteen participants (12 females, all right-handed) performed the experiment (age ranging from 21 to 45 years; $M = 25$, *95% CI* $= [22; 29]$). After pre-processing the data, we discarded three subjects:

- one subject reported "external distraction" on half experience-sampling probes (see Experience-sampling probes);
- a second subject reported 85% "task-related MW" but only one "task-unrelated MW"; moreover, only one epoch linked to "focus" was free from artifacts (two out of three epochs were discarded due to muscle activity). Subsequent questions at the end of the experiment revealed that he/she did only partially understood the difference between task-related MW and task-unrelated MW;
- a third subject displayed many movements during the experiment (foot tapping, jaw clench, arm movements), which were later found heavily decreasing data quality.

This resulted in 15 subjects in the analysis. The participants in this study were volunteers from the ONERA Company (ONERA, the French Aerospace Lab) or Marseille University. They received 20€ vouchers (cards for online payment) for the experiment. All the participants had normal or corrected-to-normal visual acuity and hearing, had no neurological or psychiatric disorders, and were not under any medication. All participants signed a written declaration of informed consent. The procedure was approved by ONERA ethical committee and was conducted in accordance with the World Medical Association Declaration of Helsinki.

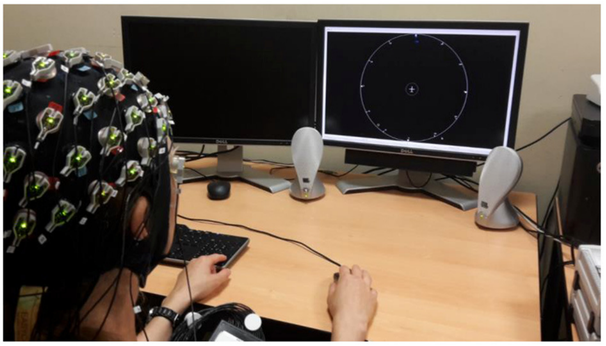


FIGURE 1 | Experimental setup. The participant is equipped with the EEG system and sits in front of the right screen (LIPS screen). Speakers are on both sides of the right screen. The left screen is used to display attentional probes.

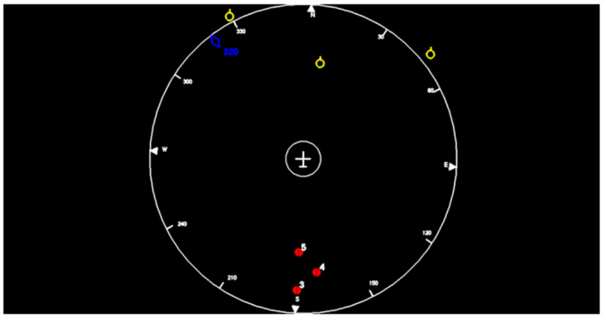


FIGURE 2 | Screenshot of the LIPS interface. The plane in the center is static and the surrounding (yellow and red numbered symbols) are moving. During the left and right avoidance maneuver, again, the plane remains static and the background rotates.

Experimental Tasks

Environment

Participants were seated in front of a desk with two screens, two speakers, a keyboard, and a mouse (see **Figure 1**). Participants performed two tasks in parallel: a visual task and an auditory task. The visual task, an obstacle avoidance task (see Visual task), was displayed on the right screen. The auditory task was presented with speakers on the left and right sides of the right screen, which sent the beeps at semi-random intervals as well as the continuous modulated brown noise (see Auditory task). On the left screen, attentional probes appeared semi-randomly (see Experience-sampling probes).

Visual Task

The visual task consisted in the supervision of an obstacle avoidance simulator displayed on the right screen (the Laboratoire d'Interactions Pilote-Système (LIPS), or Pilot-System Interactions Laboratory an ONERA distributed simulation environment). The aircraft moved at a constant speed. It was displayed in white onto a 22″ LCD monitor (with a 1,024 × 768 pixel resolution and a 60-Hz refresh rate) located about 50 cm from the participant in an unlit room.

The visual task displayed an unmanned air vehicle (UAV) depicted as a plane seen from above. The vehicle stayed at the center of a 2D radar screen (right screen, see **Figure 2**) and moved following waypoints arranged in a semi-straight line with clusters of obstacles along the way (every 45 s on average). Each cluster could contain between one and five obstacles, including one on the trajectory. When an obstacle was present on the trajectory (a situation called "conflict"), the autopilot detected it and initiated a left or right deviation, depending on the placement of the obstacles. Once the obstacle on the trajectory had been cleared, the UAV initiated another maneuver to come back on its initial straight-line trajectory. Participants were instructed to monitor the UAV, acknowledge its decisions, and correct any mistake the autopilot might make, i.e., choosing an avoidance trajectory that would result in an impact with another obstacle. In more details:

- Whenever they saw the autopilot changing the trajectory, participants clicked on an "*Acquittement*" (acknowledgment) button to acknowledge automated avoidance decisions (twice per conflict, once to acknowledge avoidance of the object and once to acknowledge the return to normal trajectory after avoiding the object);
- If they detected an incoming collision, they clicked on the button "*Changement d'altitude*" (change height) so that the UAV would perform an emergency descent to avoid colliding with the obstacle.

In both cases, a feedback message was displayed to the participants whenever they clicked.

Auditory Task

An auditory task was proposed at the same time as the visual task. Participants had to react as fast as possible to beeps (100 ms duration, 1,000 Hz frequency). Participants had 1 s to answer to these beeps presented at semi-random intervals; if they did not respond within the given time, the auditory stimulus was counted as a miss. This task was supported by E-Prime 2.0 (Psychology Software Tools, 2018). The auditory task was used to measure attention through reaction time and EEG measures (see Electroencephalography).

On top of the beeps for the auditory task, we played using E-Prime a background brown noise modulated in amplitude to elicit ASSR. Amplitude modulation was chosen as the most widely used steady-state stimuli (Picton et al., 2003) better tolerated by people than clicks (Voicikas et al., 2016). We first generated brown noise using the *acoustics.generator.brown* function (felipeacsi and Rietdijk, 2018). This signal was then modulated with a 50% and 40-Hz sinusoidal amplitude modulation. Because E-Prime loads file sounds as the experiment develops, a 1-h file would have exceeded the cache memory. To allow for easier loading, we divided the sound into 5-s soundtracks played one after the other in a loop (**Supplementary Audios 1–3**). To avoid participants to develop explicit or implicit learning with repetitive sound features, we generated three different 5-s soundtracks, which E-Prime played in random order. Tests before the experiment did not reveal any audible problem when switching between soundtracks, nor did participants realize it when asked after the experiment (Agus and Pressnitzer, 2013). We used Python 3.6 to generate modulated background brown noise with the base packages *acoustics*, *wave*, *math*, and *random* (Python Software Foundation, 2018).

Experience-Sampling Probes

On average, every 2 min, an experience-sampling probe programmed with E-Prime 2.0 (Psychology Software Tools, 2018) appeared on the left screen (**Figure 1**). For technical reasons, the visual task (obstacle-avoidance task) was not paused when the experience sampling probes appeared. Participants were asked to answer the probe as soon as it appeared, and any successful or failed trial on the obstacle-avoidance task during this interval was not taken into account to compute their performances on the visual task. Participants were informed that the questionnaire probes were for informational purposes only and were not used to assess performance.

Participants were required to answer the following question (originally in French): "When this questionnaire appeared, where was your attention directed?" Answers could be "On the task" (focused, e.g., thinking about the next obstacle, the decision to make, the incoming waypoint), "Something related to the task" (task-related MW, e.g., thinking about performance, interface items, last trial), "Something unrelated to the task" (task-unrelated MW, e.g., thinking about a memory, their last meal, or a body sensation) or "External distraction" (e.g., conversation, noise). The preceding examples were given to participants to illustrate each category before the experiment. We were primarily interested in reports of being focused or having task-related or task-unrelated MW. The possibility of reporting "task-related MW" was proposed to avoid participants reporting task-unrelated MW when thinking about their performance (Head and Helton, 2016). The answer "External distraction" was proposed to avoid participants reporting MW if they were distracted by a signal external to themselves and the task.

Procedure

Sessions started with an explanation of the two tasks, followed by a 10-min training period and a 55-min session. During this study, participants had to perform the visual task (supervise the UAV avoiding obstacles and acknowledge or correct any mistake, see Visual task) and the auditory task (press a button as fast as possible when hearing a beep, see Auditory task) at the same time. The session contained 70 clusters of obstacles for a total of 210 obstacles. Clusters were separated by 45 s on average. All autopilot decisions and collisions were predefined and, therefore, they were the same for all subjects. The autopilot made two errors initially placed randomly (3% errors; errors on trials 31 and 52 for all subjects). This low error rate was chosen to have a relatively safe system and reproduce ecological OOTL conditions.

Parallel to the visual task, participants performed the auditory task and had to react to infrequent beeps by pushing "Enter" button as fast as possible with their left hand. This secondary task served as a way to measure attention (see Measures and analysis for the exact measures reported). They were explicitly told that beeps and experience-sampling probes were to be treated as fast as possible, whatever was happening on the obstacle-avoidance task. Beeps were presented every 20–40 s. On average, one out of three beeps was followed by an attentional probe. In total, 32 probes were displayed during the whole session. The distribution of the experience-sampling probes was not correlated with events on the obstacle-avoidance task, to minimize performance influence on experience-sampling reports. We instructed participants not to pay attention to the ASSR background sound.

Measures and Analysis

We used R-Studio 1.1.456, R 3.5.1 (RStudio Team, 2015; R Core Team, 2016) for statistical analysis, and Matlab 2018a (The Mathworks Inc., 1992), EEGLAB (Delorme and Makeig, 2004), and FieldTrip (Oostenveld et al., 2010) to filter and analyze EEG data. All 95% CIs reported hereafter were computed using the *boot* R package with 10,000 iterations with normal bootstrap approximation (Canty and Ripley, 2017).

All linear mixed-effect analyses used the R *lme* function to create the models (Bates et al., 2017), with a random intercept for subjects to account for our repeated-measure design. Each time, we visually inspected residual plots to spot any obvious deviations from normality or homoscedasticity. We assessed the influence of predictors by creating a baseline model and then added each predictor in turn; we compared each model with the previous one to verify if adding a predictor significantly reduced uncertainty. The R *Anova* function was used to compare models by performing likelihood-ratio tests between given models and report the χ^2 value (R Core Team, 2016). We chose type 2 sum of squares or type 3 sum of squares when there were interactions to consider between predictors. *Post hoc* tests were conducted using the *glht* and *mes* functions on the complete model (R Core Team, 2016).

Subjective Measures

Subjective measures consisted of the answers to the experience-sampling probes. We split the 55-min sessions into four blocks of ∼14 min containing eight experience-sampling probes each. We focused on task-related and task-unrelated MW frequency evolution over time and conditions using linear mixed-effect analysis. We considered blocks as a four-level categorical variable. Without specific a priori predictions regarding the block-by-block evolution, we conducted Tukey's *post-hoc* tests on the complete model.

Behavioral Measures

To assess performance in the auditory condition, we recorded accuracy and reaction time related to beep answers (the difference between start of the beep and the button press). The influence of attentional states and blocks on reaction time was analyzed using a linear mixed-effect analysis. We conducted Tukey's *post-hoc* tests to break the potential effects of blocks.

Electroencephalography

We used the ActiCHamp system and Brain Vision software (Brain Products, 2018) to record scalp potentials. A total of 64 Ag–Cl electrodes were mounted on a standard elastic cap at the standard sites of the 10–10 International system (Oostenveld and Praamstra, 2001). Impedance was kept below 5 kΩ for all electrodes. The Fpz electrode was used as the ground electrode. We used electrooculographic sites to capture eye movements. We chose the left mastoid FT9 electrode as a reference for recording.

We were interested in the influence of attentional states on stimuli perception and treatment. Beeps served as a way to measure attention through ERPs. We selected N1 (a marker of perception) and P3 (a marker of stimuli processing) elicited by the auditory task. Following the literature, we analyzed the 180–200 ms interval average on electrodes Fz, Pz, and Cz for the P3 and N1 components (Kam et al., 2011, 2014; Kam and Handy, 2013). Similarly, we chose the 380–420 ms interval average and the same electrodes for P3 component.

Regarding spectral analysis, we also used the auditory task and the time immediately preceding beeps. We focused on the upper alpha band because previous studies repeatedly revealed consistent results for the lower and upper alpha band (e.g., Benedek et al., 2011; Jauk et al., 2012). We also investigated the ASSR frequency. We chose the electrodes Pz, P1/2, P3/4, P5/6, POz, PO3/4, Oz, and O1/2 for alpha to cover the parieto-occipital region. Previous studies observed higher alpha amplitude linked with visual sensory inhibition in this region, in line with the MW perceptual decoupling (Foxe et al., 1998; O'Connell et al., 2009; Benedek et al., 2014). For the ASSR, we monitored the 39.5–40.5 Hz band where the stimulus was supposed to elicit a peak. We used the sites FCz, FC1/2 for ASSR, which had already been used by Saupe and colleagues in experiments investigating ASSR and attention (Saupe et al., 2009b; Keitel et al., 2011).

Each time an experience-sampling probe appeared, a signal was sent to the ActiCHamp software to record a trigger on the EEG signal. Similarly, another trigger was sent when participants answered the probe, whose value depended on attentional state reported by participants, and a last signal was sent by the auditory task whenever a beep played. Triggers sent by beeps served as a synchronization point to study EEG metrics, whereas triggers of probes served to classify the attentional state of participants when the beep immediately preceding played. The timing of the overall setup was tested and revealed no important deviations. We used Matlab, EEGLAB, and FieldTrip to import, re-reference, filter, epoch, remove ICA components, and build our design. The exact filtering pipeline was as follows:

- Add coordinates to existing 63 electrodes using template 10–20 location (BESA spherical format; function used: *pop_chanedit*).
- Re-reference data to FT9 and FT10 channels (Yao et al., 2005; Griskova et al., 2007; Kam et al., 2011, 2012; function used: pop_reref).
- Filter using a two-pass pass-band Butterworth filter to avoid shifting introducing the signal. The pass-band used was [0.01; 30] Hz for ERPs and [0.01; 100] Hz for ASSR and alpha (function used: *ft_preprocessing*).
- Interpolate electrodes when the line noise was deemed too important: if it displayed (1) variation above ∼300 μV amplitude, (2) variation uncorrelated to other electrodes around it, and (3) previously mentioned issues were spotted on at most one subject, as the same problem found on multiple subjects would mean that the electrode itself is faulty and should be suppressed from the overall study (overall decision made after visual inspection; on average 0.1 electrode interpolated per participant for ERPs and 0.5 electrodes interpolated per participant for ASSR and alpha wave).
- Create epochs by taking signal intervals around beeps. The interval was [−800; 1,000] ms for ERPs and [−5,000; 0] ms for ASSR and alpha (on average 31.7 epochs per participant; function used: *pop_epoch*).
- Remove the baseline of each epoch: for ERPs, we took the average signal in [−200; 0] ms and subtract it from the whole epoch; for ASSR and alpha, remove base power of each frequency (function used: *pop_rmbase*).
- Discard epochs when they were heavily contaminated by muscle artifacts which would lower ICA power (decision made after visual inspection, although multiple backs and forth were made to determine ICA impact tolerance; on average 3.3 epochs discarded per participant for ERPs, 4.4 epochs discarded per participant for ASSR and alpha wave; function used: *eegplot*).
- Run the ICA with option "extended, 1" also reducing the number of dimension by one due the rank deficient matrix (function used: *pop_runica*).
- Discard components in case of ocular movements (high power coupled with activity frontal, dissymmetrical from both eyes perspective, spatially and temporarily narrowed), blinks (high power coupled with activity frontal, symmetrical from both eyes' perspective, spatially and temporarily narrowed), other muscle activity (very high power coupled with activity spatially and temporarily narrowed), and electrode malfunction (very high power, activity centered on one specific electrode). The final decision was made after visual inspection (no epochs discarded for ERPs, on average 1.6 epochs discarded per participant for ASSR and alpha wave; function used: *pop_selectcomps*).

We then exported data to R to perform statistical analysis. We used a linear mixed-effect analysis to look at the influence of attentional states on ERPs, alpha, and ASSR amplitude.

RESULTS

MW Frequency Analysis

Participants reported on average 31.3% task-related MW (*SD* = 4.4%) and 36.6% task-unrelated MW (*SD* = 5.0%, see **Figure 3**, **Supplementary Data Sheet 1**). This rate is consistent with previous studies (Smallwood et al., 2006; Smallwood and Schooler, 2015; Gouraud et al., 2018a,b). Each participant reported on average 1.5% "External distraction" reports (*SD* = 1.21). Considering this low rate, we discarded "External distraction" reports and adopted the ternary approximation of attentional states (i.e., either focused, task-related MW, or task-unrelated MW). All participants answered all 32 probes, except one participant who did not answer four probes.

Blocks did not significantly influence task-related MW. On the contrary, blocks significantly influenced task-unrelated MW rates, $\chi^2 = 12.13$, $p = 0.007$. *Post-hoc* tests revealed that

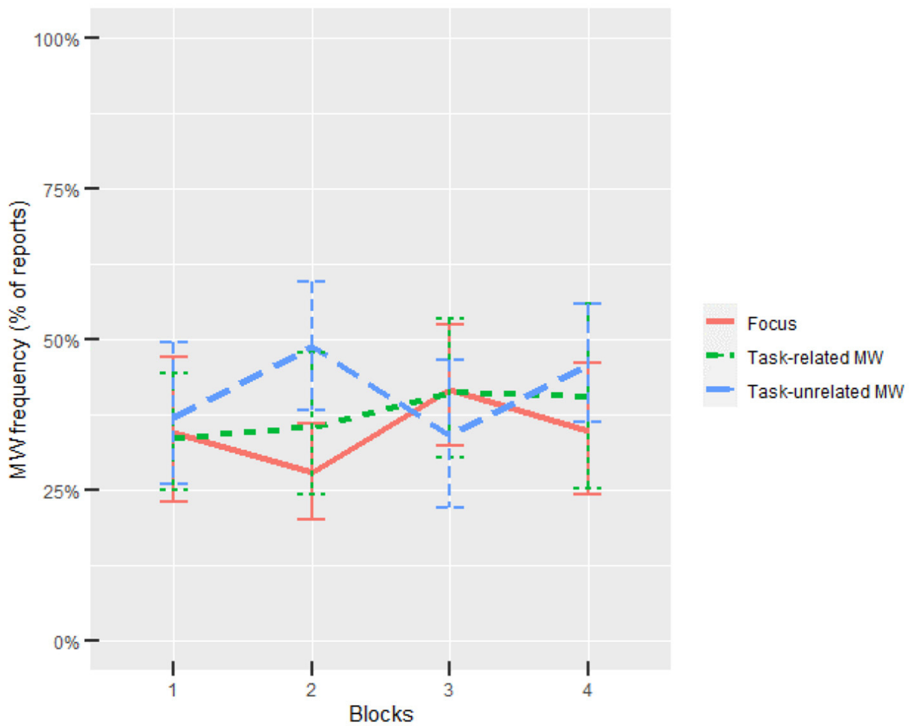


FIGURE 3 | Task-related and task-unrelated MW evolution through blocks. Error bars show the 95% CIs based on bootstrap.

TABLE 1 | Influence of blocks on task-related and unrelated MW frequency.

		Task-related MW		Task-unrelated MW	
Effect added	***df***	χ^2	***p*-value**	χ^2	***p*-value**
Block	3	0.30	0.828	**12.13**	**0.007**

Bold values are significant results.

TABLE 2 | Influence of attentional states and blocks on beep reaction time.

Effect added	*df*	χ^2	*p*-value
Attentional states	2	2.89	0.24
Block	3	**25.52**	**<0.001**
Attentional states: blocks	6	10.09	0.121

Bold values are significant results.

task-unrelated MW rate were significantly higher under the second block compared with the first and third blocks, $p = 0.021$, $d = 0.55$, $p = 0.010$, $d = 0.62$, respectively. All results from model comparisons are gathered in **Table 1**, bold values being significant.

Auditory Task: Reaction Time to Beeps

The auditory task performance was investigated using reaction time when presented a beep followed by a probe. Participants reacted to on average 31.3 beeps out of the 32 presented. Attentional states did not influence reaction time. On the contrary, there was a significant influence of blocks on reaction time, $\chi^2 = 25.52$, $p < 0.001$. *Post-hoc* tests revealed that participants were significantly slower during the fourth block compared with the first and third blocks, respectively ($p = 0.007$, $d = 0.48$ and $p = 0.016$, $d = 0.28$). All results from model comparisons are gathered in **Table 2** and illustrated in **Figure 4**.

Auditory Task: Influence of Attentional States on ERPs

The amplitude evolution of ERPs elicited by the auditory task (beeps) was investigated. Attentional states significantly influenced both N1 and P3 components (see **Table 3** and **Figure 5**). *Post-hoc* tests revealed that for the N1 component, reports of task-unrelated MW were accompanied with a lower amplitude ($M = -6.06$ μV, *95% CI* = [−8.01; −4.12] μV) compared with periods of focus ($M = -9.39$ μV, *95% CI* = [−12.21; −6.60] μV), $p = 0.024$, $d = 0.36$. For the P3 component, the statistics showed a significantly higher amplitude for task-related MW ($M = 12.69$ μV, *95% CI* = [9.28; 16.13] μV) compared with focus periods ($M = 8.20$ μV, *95% CI* = [5.54; 10.85] μV), $p = 0.009$, $d = 0.16$.

Visual Task: Influence of Attentional States on Alpha Wave Amplitude

Alpha wave power evolution before experience-sampling probes was investigated. Results showed a significant influence of attentional states on alpha amplitude (see **Figure 6** and **Table 4**, bold values being significant), $\chi^2 = 8.35$, $p = 0.015$. *Post-hoc*

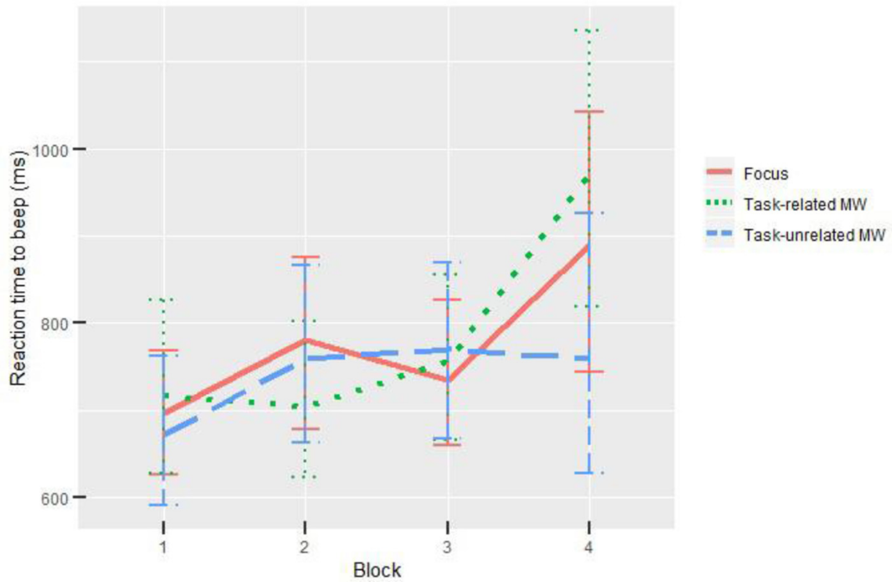


FIGURE 4 | Influence of blocks and attentional states on beep reaction time. Error bars show the 95% CIs based on bootstrap.

TABLE 3 | Influence of attentional states on the amplitude of the ERP components N1 and P3.

Effect added	*df*	N1 component		P3 component	
		χ^2	*p*-value	χ^2	*p*-value
Attentional states	2	**9.41**	**0.009**	**8.83**	**0.012**

Bold values are significant results.

tests showed significantly higher alpha amplitude during task-unrelated MW ($M = 53.83\ \mu V^2/Hz$, *95% CI* = [52.35; 55.31] $\mu V^2/Hz$) compared with focus episodes ($M = 53.03\ \mu V^2/Hz$, *95% CI* = [51.90; 54.16] $\mu V^2/Hz$), $p = 0.014$, $d = 0.27$. Other comparisons (task-related MW vs. focus, task-related MW vs. task-unrelated MW) were not significant.

Influence of Attentional States on ASSR Amplitude

No influence of attentional states on ASSR amplitude was uncovered (**Figure 7**). However, spectral plots still revealed a peak at 40 Hz, showing that the ASSR was visible on participants' spectrum even during this complex task (see **Figures 8**, **9**). Should anyone want to reuse this background noise for other ASSR activities within aeronautical-inspired environments, we mention that 12 participants out of 18 reported that they felt the noise was similar to a propeller airplane.

DISCUSSION

The aim of this study was to evaluate the viability of MW neuronal markers in complex ecological automated environments, and to help characterize features of the attentional decoupling in these settings. We chose an automated obstacle avoidance task that participants had to supervise while reacting as fast as possible to beeps they heard. EEG signal was chosen to acquire cerebral activity in the form of ERPs, alpha wave amplitude, and ASSR. To yield detailed results, we decomposed MW into task-related and task-unrelated acquired using attentional probes. We decomposed the 40-min task into 4 blocks of 10 min each. Participants did not show any increase in task-related or non–task-related MW during the time spent on the task although more task-unrelated MW emerged during the second block. When analyzing ERP components created by beeps, we observed lower N1 component amplitude during task-unrelated MW, while P3 component had higher amplitude during task-related MW, compared with other attentional states. Alpha wave activity was higher in parieto-occipital regions during task-unrelated MW compared with other attentional states. Finally, ASSR was clearly elicited, but its amplitude was not significantly influenced by attentional states. Overall, these results underline the complex influence of the MW perceptual decoupling on operator's behavior in ecological environments and have several implications when considered together.

Measuring the Impact of MW

Taken together, the observed effects support a reduction in cortical processing of the external environment during task-unrelated MW. First, for the auditory task, N1 component elicited by the beeps had a lower amplitude during task-unrelated MW, indicating a state of reduced perception of stimuli already identified by Kam et al. (2011). Participants who experienced task-unrelated MW were less receptive to the beeps. Nevertheless, only a non-significant trend could be observed in reaction times (**Figure 4**), with subjects being faster during the fourth block

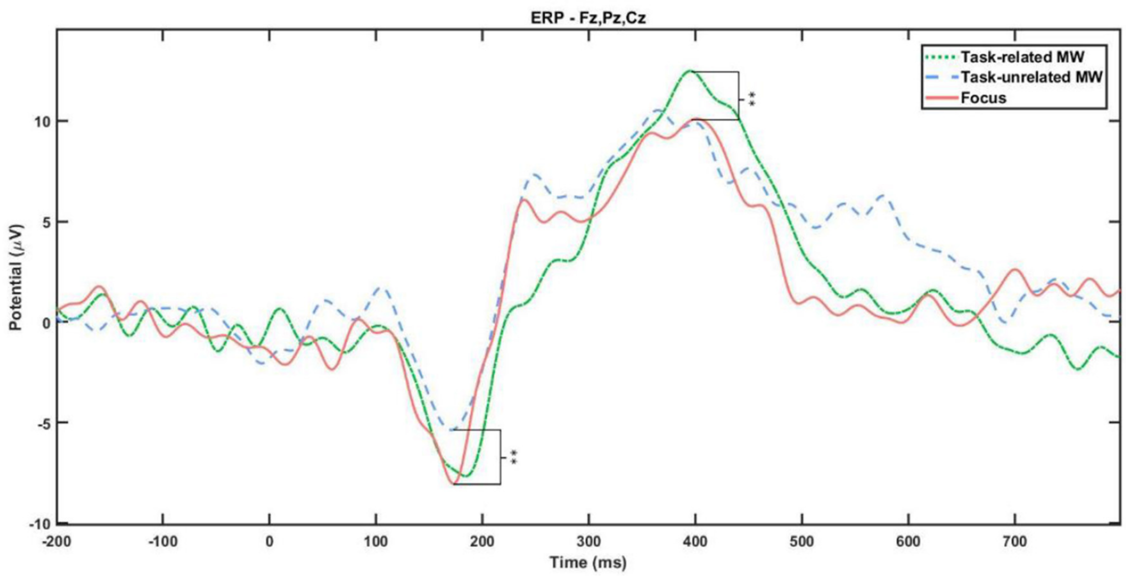


FIGURE 5 | Beep ERP signal for task-related MW (green), task-unrelated MW (blue), and focus (red) attentional states.

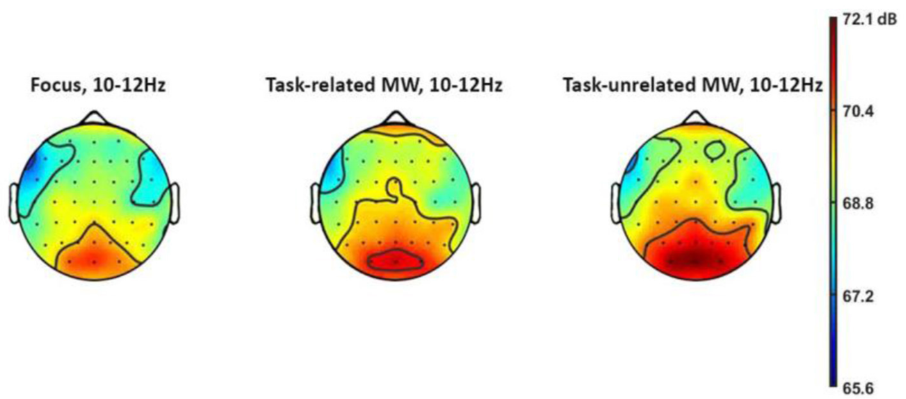


FIGURE 6 | Topography of alpha frequency for each attentional state.

for task-unrelated MW compared with other attentional states. Subjects may have focused, maybe even attention-tunneling, on the visual task when being focused or in task-related MW. On the contrary, being in task-unrelated MW may have led participants to use strategies favoring speed over precision, without significant impact on the accuracy due to the low difficulty of the task (Salomone et al., 2021).

Second, regarding the visual task, the increase in alpha power in the parieto-occipital lobe shows that participants inhibited visual perception during MW episodes (Foxe and Snyder, 2011; Benedek et al., 2014; Clayton et al., 2015). Although the debate still exists on alpha power, both analyses are congruent and consistent with research sharing the same features, i.e., probe-caught MW (Baird et al., 2014), visual (Compton et al., 2019), and ecological task (Baldwin et al., 2017). MW creates a decoupling from the task at hand, even in complex bimodal environments. Our results are a first step toward filling the gap between real consequences of MW (Galera et al., 2012; Berthié et al., 2015) and EEG research in laboratory settings (Kam, 2010; Kam et al., 2019). Taken together, visual and auditory analyses support the multimodal influence of MW in complex environments (Kam et al., 2011), although our setup does not allow us to make quantified claims and compare modalities.

We observed no effect of attention on ASSR amplitude, even though its evoked power was visible on the EEG spectrum of the

TABLE 4 | Influence of attentional states on alpha and ASSR amplitude.

Effect added	*df*	Alpha power (log)		ASSR amplitude	
		χ^2	*p*-value	χ^2	*p*-value
Attentional states	2	**8.35**	**0.015**	2.55	0.279

Bold values are significant results.

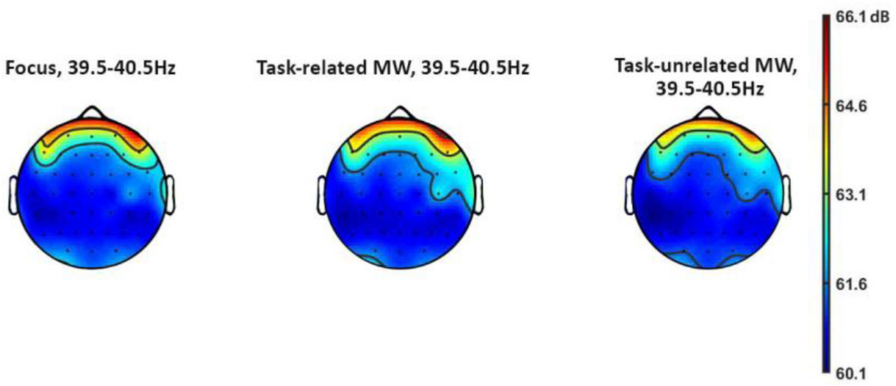


FIGURE 7 | Topography of ASSR frequency for each attentional state.

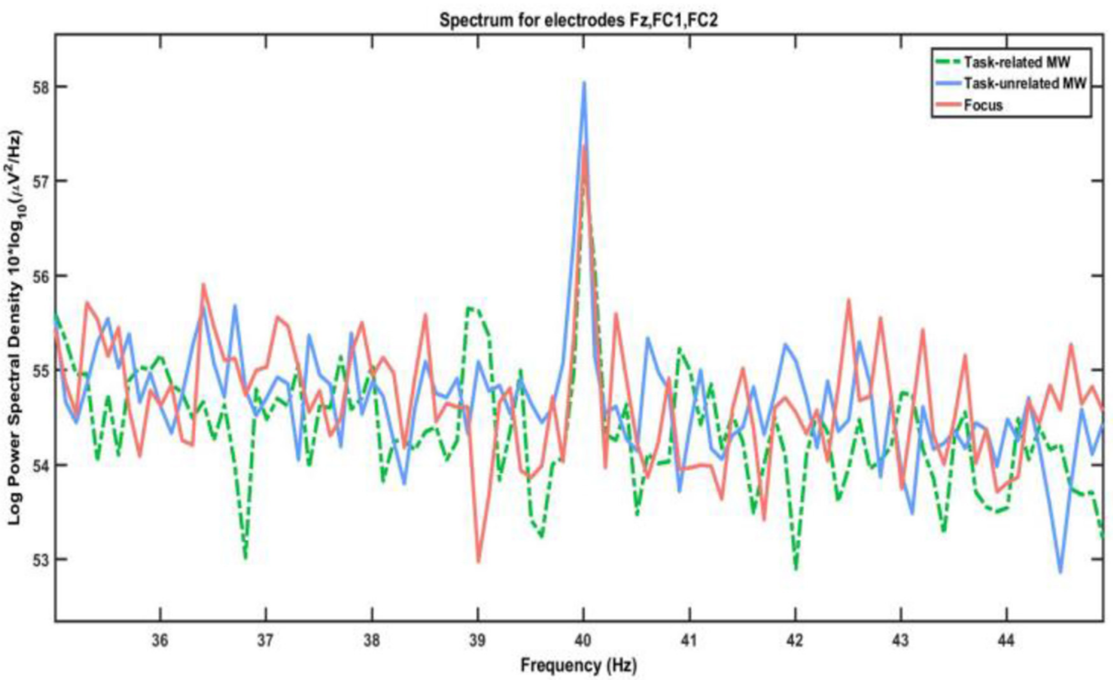


FIGURE 8 | Spectrum of 35–45 Hz interval for each attentional state.

participants. This outcome is in line with the results of O'Connell et al. (2009) regarding the absence of amplitude modulation of MW on SSR. It is possible that our experiment did not succeed because of its features, such as the use of amplitude modulation instead of clicks (Voicikas et al., 2016) or the insufficient number of participants. Another possibility may be that SSR produced by non-target background noise is already being reduced by participants instructed to ignore it from the start; it may therefore not be further influenced by MW. However, this hypothesis is in contradiction with both literature on ASSR in attention modulation settings (Skosnik et al., 2007; Müller et al., 2009; Mahajan et al., 2014) and our own results regarding lower N1 amplitude during task-unrelated MW. To account for this observation, a final explanation may be that internally directed attention like MW is fundamentally different from the evolution of external direction between sensory modalities. In this case, the absence of amplitude modulation would show that MW does not impact the earliest stages of perception, allowing for a basic processing of external stimuli. Further work in this area is needed to provide robust conclusions.

Gradual Impact of MW

Important differences were highlighted between task-related and task-unrelated MW, supporting the existence of "depth" or

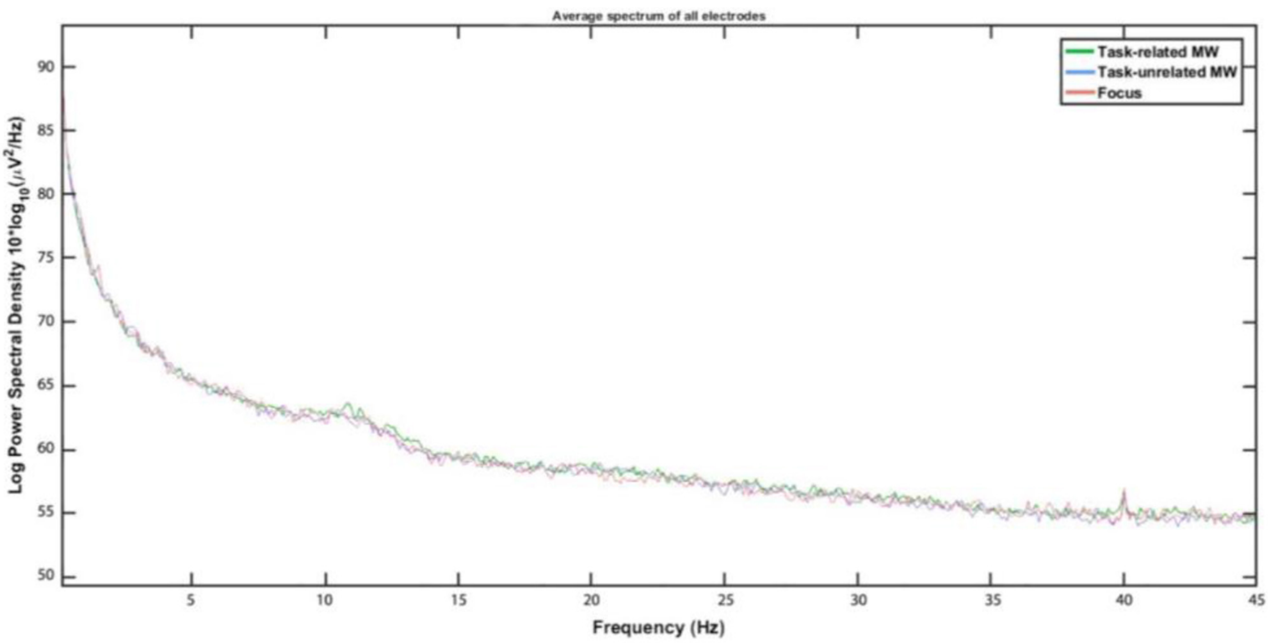


FIGURE 9 | Spectrum of 0–45 Hz interval for each attentional state.

"intensity" (related to the decoupling) in MW episodes. During task-unrelated MW, participants inhibited perception of auditory stimuli (as shown by the N1 amplitude), but not during task-related MW compared with focus moments. On the contrary, auditory information processing (P3 amplitude) was higher during task-related MW than during focus intervals. Participants reporting being focused may actually focus on the visual part of the task (the most cognitively demanding) while inhibiting all auditory stimuli, whether relevant to the task or not. On the other hand, task-related MW may create a more superficial decoupling than task-unrelated MW. This mental state may redirect attentional resources from the exhausting visual task to listening to auditory cues, thus participating in a more balanced resource allocation independently of task demand. Unfortunately, we did not observe differences in performance, i.e., reaction time during the auditory task. It is likely that because the processing of auditory stimuli did not require much cognitive resources, superficial perception was enough to perform it.

Previous explanation remains very conditional, as the available observations are not sufficient to definitely establish the depth of MW. A graded MW with a different decoupling could explain why we are most of the time able to perform tasks while being in MW, while sometimes we make clear errors that could have been avoided with our full attention (Cheyne et al., 2006; Carriere et al., 2008; Farley et al., 2013). Two protocols may complete the present study in relation to MW depth: using the same experiment, but asking the participant to ignore the beeps; the irrelevance of beeps may thwart interesting results when analyzing the influence of task-related MW on ERPs. Another possibility would be to use the same experiment once again, but this time participants would have two different beeps to react to, each associated with a different button. The needs for more processing of auditory stimuli could link the performance data to MW decoupling depth. Nevertheless, more data are needed to rule over the depth dimension.

Factors Stimulating MW Emergence

In this experiment, MW rates remained mostly stable through time-on-task, only the second block exhibiting higher task-unrelated MW rates compared with the first and third ones. We witnessed similar behavior in our previous study (although here MW increased in the middle of the task instead of decreasing, see Gouraud et al., 2018a). Literature generally agrees that MW rates should increase with time-on-task (Smallwood et al., 2002; Pattyn et al., 2008; Risko et al., 2012; Gouraud et al., 2018a) although several studies failed to observe such behavior (Thomson et al., 2014; Arnau et al., 2020). Nevertheless, the exact link between MW and time-on-task may be mediated by task difficulty, i.e., task demands in attentional resources (McVay and Kane, 2009; Krimsky et al., 2017). We have already used as the only task our automated UAV monitoring environment in previous experiments without observing more MW, which shows that the multitasking did not require much attention from participants (see Mind Wandering Frequency Analysis and Gouraud et al., 2018a,b). Moreover, attention demand remained constant throughout the task, which further decreased the possibility of bias in our subsequent analysis. To explain the lack of increase in MW with time-on-task, a first explanation might be that participants, aware of the overall duration of the experience, sensed time passing by and reengaged in the task in the second half (Arnau et al., 2020). The lack of MW increase with time-on-task might also be due to automation errors, placed at the ends of the second and third blocks. A third possibility might be explained by a too disruptive setup (e.g., beeps allowing reengagement, EEG being too uncomfortable). However, our previous experiments with the same visual environment, but no auditory stimuli, yielded equivalent attentional state percentages on average (Gouraud et al., 2018a,b).

More generally, the question of what conditions will stimulate the emergence of MW remains, both in experiments and in the open. Time-on-task plays an important role (Smallwood et al., 2002). However, it may not be the only factor: on top of various individual features linked with different MW rates [training in Casner and Schooler (2015); positivity in Hancock (2013); gender in Mar et al. (2012); creativity in Zedelius and Schooler (2016)], the very nature of tasks to perform could influence MW and its evolution. In particular, operators faced with increased automation see their relation to the task dramatically modified. We already investigated the influence of automation levels in a previous experiment (Gouraud et al., 2018b) without significant differences in MW rates between a manual and an automated condition.

Nevertheless, many dimensions of automation that could influence MW rates remain unexplored. One of the main impacts of higher automation is a drop in operators' sense of control or agency (Haggard, 2017). Sense of agency is the experience of identifying oneself as the author of an action and its consequences (Gallagher, 2000). This form of self-awareness is important not only for motor control but also for causal responsibility and serves as a key motivational force for human behavior. Recently, it has been shown that the sense of agency could be dramatically impaired when interacting with automation (Berberian, 2019). While co-workers develop a form of we-agency (Crivelli and Balconi, 2010; Obhi and Hall, 2011), the same does not stand true for human–system cooperation (Wohlschläger et al., 2003a,b; Glasauer et al., 2010; Sahaï et al., 2017). Similarly, there is a loss of agency when operators' tasks shift from working a system to monitoring it (Berberian et al., 2012). Even though automation generally brought safer and more productive systems, the loss of agency could generate task disengagement and be one of the main reasons why operators are unable to regain manual control in critical situations (Bainbridge, 1983; Endsley and Kiris, 1995; Cummings, 2004; Louw et al., 2015b; Berberian et al., 2017). Critically, Wen and Haggard (2018) have highlighted important differences in attention allocation correlated with differences in the sense of agency: the loss of a sense of control could decrease the allocation of attentional resources to stimuli relevant to the task at hand. In this context, loss of agency may have a significant influence on MW rates. To our knowledge, no experiment has investigated the relation between MW and agency.

MW and Operator Engagement Issue

As our results showed, distinguishing different types of MW revealed different impacts on EEG measures, while the absence of MW influence on ASSR may highlight a fundamental difference between internally and externally directed attention. Despite these unknowns, our results add to the existing literature supporting the decoupling hypothesis and linking MW to a form of attentional disengagement. Indeed, task engagement strongly modulates performance through goals and motivation (Bedny and Karwowski, 2004; Fairclough et al., 2013; Leontiev, 2014), concepts that are strongly linked with MW (Cheyne et al., 2009; Danckert, 2017; Gouraud et al., 2018b). MW could exacerbate task disengagement by highlighting the discrepancy between entertaining thoughts and the ungratifying present (Smallwood and Schooler, 2006; Eastwood et al., 2012) and drawing attention to one's own failure to maintain vigilance (Critcher and Gilovich, 2010; Westgate and Wilson, 2018). Other researchers believe that MW may be just a symptom of boredom: internal sources of stimulation could serve as a second-best option when external tasks fail to keep us focused (Singer, 1975; Bench and Lench, 2013). Neurologically, MW episodes are characterized by the deactivation of the dorsolateral prefrontal cortex (DLPFC, see Christoff et al., 2009; Stawarczyk et al., 2011). DLPFC interacts with dorsal and ventral attentional pathways to shift and focus attention on the most relevant stream of task-related information (Johnson and Zatorre, 2006). It is a network thought to play a crucial role in maintaining task engagement (Curtis and D'Esposito, 2003). MW is thought to represent the lower end of a continuum of task engagement (Lee, 2014; Dehais et al., 2020).

MW pertains to a wider collection of mental states linked to engagement and negatively impacting performance. These suboptimal neurocognitive states are investigated by neuroergonomics, whose purpose is the study of the human brain in relation to performance at work and in everyday settings (Parasuraman, 2011; Gramann et al., 2017). The development of this field has been facilitated by the twenty-first century revolution in our understanding of neural mechanisms, but also by recent developments in advanced and portable neuroimaging techniques (Dehais et al., 2020). Several attempts have been made to identify MW features within dry EEG signals, with success on ERPs and alpha waves (van der Wal and Irrmischer, 2015; Kam et al., 2019). Functional Neuro InfraRed Spectroscopy (fNIRS) has also demonstrated its capability to detect MW episodes in ecological simulation by monitoring the Default Mode Network (Durantin et al., 2015), a network involved in attention drifting processes (Raichle et al., 2001; Konishi et al., 2015; Golchert et al., 2016). Both dry EEG and fNIRS could be integrated into operational environments with little disruption for the user (Mullen et al., 2015; OpenBCI, 2016; This Place, 2016; SmartCap, 2020). On top of neuroimaging techniques, oculometry has also been substantially improved over the past decade, producing efficient, small, and cheap devices. It has demonstrated a high sensitivity to MW in safety-critical environments, although only in simulators (Louw et al., 2015a; Louw and Merat, 2017). Thanks to these systems and models, neuroergonomics could help translate MW findings from psychology and neurosciences into procedures changes to enhance safety in the industry.

CONCLUSION

We presented the results of an EEG study with a visual (monitoring and correction of an automated UAV avoiding obstacles) and an auditory (infrequent beep which required fast button press) task presented simultaneously with the aim to understand the cerebral signature of MW. Participants also heard a background noise designed to elicit ASSR. We saw that task-related and task-unrelated MW exhibit a different EEG signature, whether it is on ERP components or on alpha waves, suggesting the existence of depth in perceptual decoupling. Our results also stress the need to carefully discriminate MW dimensions when evaluating MW-induced decoupling. Finally, the absence of MW hallmark on ASSR amplitude does not support the possibility to use SSR to study MW continuously. However, it also means that the earliest stages of perception may not be impacted by attentional decoupling.

Overall, our results highlight the crucial need to study the neural correlates of MW to identify its exact influence on operators. Even though the setup involved remained highly controlled and laboratory related, our tasks were relatively close to complex automated environments encountered in operations, and more specifically teleoperations. Contrary to recent claims (Neigel et al., 2019), MW pervasive effects have been widely reported in monotonous ecological simulations (He et al., 2011; Casner and Schooler, 2014, 2015; Louw et al., 2015a,b; Baldwin et al., 2017; Gouraud et al., 2018a,b) and real environments (Galera et al., 2012; Berthié et al., 2015). Moreover, they are perfectly integrated in several recent neuroscientific models (Pattyn et al., 2008; Dehais et al., 2020). Other problems teleoperations should overcome involve operators' ability to mentally jump into a situation while being physically away and should be specifically assessed, and the related issues studied. Distraction and other forms of inattention are already a significant safety problem within the transport industry, e.g., in the air (Loukopoulos and Field, 2001; Casner and Schooler, 2015) or on the road (Galera et al., 2012; Berthié et al., 2015). In this context, a better understanding of MW, which participates in operator distraction, is crucial to limit distraction consequences. It is essential that research investigates the effects of the different characteristics of MW, while the possibilities to mitigate its consequences must also be examined through both ecological setup and operational environments and the outcomes adopted by the industry. Taking the problem into account when designing the technology (Nielsen et al., 2007; Hosseini and Lienkamp, 2016) could enhance teleoperations and install it as the next important step toward full automation. In this context, neuroergonomics could bring a new perspective on this kind of suboptimal neurocognitive state to go further than broad metaphorical concepts.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by ONERA ethical committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnrgo.2021.625343/full#supplementary-material

REFERENCES

Agus, T. R., and Pressnitzer, D. (2013). The detection of repetitions in noise before and after perceptual learning. *J. Acoust. Soc. Am.* 134, 464–473. doi: 10.1121/1.4807641

Arnau, S., Löffler, C., Rummel, J., Hagemann, D., Wascher, E., and Schubert, A.-L. (2020). Inter-trial alpha power indicates mind wandering. *Psychophysiology*. 57:e13581. doi: 10.1111/psyp.13581

Atchley, R., Klee, D., and Oken, B. (2017). EEG frequency changes prior to making errors in an easy stroop task. *Front. Hum. Neurosci.* 11:521. doi: 10.3389/fnhum.2017.00521

Badue, C., Guidolini, R., Carneiro, R. V., Azevedo, P., Cardoso, V. B., Forechi, A., et al. (2019). Self-driving cars: a survey. *arXiv:1901.04407 [cs]*. Available online at: http://arxiv.org/abs/1901.04407 (accessed February 8, 2021).

Bainbridge, L. (1983). Ironies of automation. *Automatica* 19, 775–779. doi: 10.1016/0005-1098(83)90046-8

Baird, B., Smallwood, J., Lutz, A., and Schooler, J. W. (2014). The decoupled mind: mind-wandering disrupts cortical phase-locking to perceptual events. *J. Cogn. Neurosci.* 26, 2596–2607. doi: 10.1162/jocn_a_00656

Baldwin, C. L., Roberts, D. M., Barragan, D., Lee, J. D., Lerner, N., and Higgins, J. S. (2017). Detecting and quantifying mind wandering during simulated driving. *Front. Hum. Neurosci.* 11:406. doi: 10.3389/fnhum.2017.00406

Barron, E., Riby, L. M., Greer, J., and Smallwood, J. (2011). Absorbed in thought: the effect of mind wandering on the processing of relevant and irrelevant events. *Psychol. Sci.* 22, 596–601. doi: 10.1177/0956797611404083

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2017). *lme4: Linear Mixed-Effects Models using "Eigen" and S4*. R-package, v1.1-15. Available online at: https://CRAN.R-project.org/package=lme4 (accessed February 28, 2018).

Bedny, G. Z., and Karwowski, W. (2004). Activity theory as a basis for the study of work. *Ergonomics* 47, 134–153. doi: 10.1080/00140130310001617921

Bench, S., and Lench, H. (2013). On the function of boredom. *Behav. Sci.* 3, 459–472. doi: 10.3390/bs3030459

Benedek, M., Bergner, S., Könen, T., Fink, A., and Neubauer, A. C. (2011). EEG alpha synchronization is related to top-down processing in convergent and divergent thinking. *Neuropsychologia* 49, 3505–3511. doi: 10.1016/j.neuropsychologia.2011.09.004

Benedek, M., Schickel, R. J., Jauk, E., Fink, A., and Neubauer, A. C. (2014). Alpha power increases in right parietal cortex reflects focused internal attention. *Neuropsychologia* 56, 393–400. doi: 10.1016/j.neuropsychologia.2014.02.010

Berberian, B. (2019). Man-machine teaming: a problem of agency. *IFAC-PapersOnLine* 51, 118–123. doi: 10.1016/j.ifacol.2019.01.049

Berberian, B., Sarrazin, J.-C., Le Blaye, P., and Haggard, P. (2012). Automation technology and sense of control: a window on human agency. *PLoS ONE* 7:e34075. doi: 10.1371/journal.pone.0034075

Berberian, B., Somon, B., Sahaï, A., and Gouraud, J. (2017). The out-of-the-loop Brain: a neuroergonomic approach of the human automation interaction. *Annu. Rev. Control.* 44, 303–315. doi: 10.1016/j.arcontrol.2017.09.010

Berthié, G., Lemercier, C., Paubel, P.-V., Cour, M., Fort, A., Galéra, C., et al. (2015). The restless mind while driving: drivers' thoughts behind the wheel. *Accid. Anal. Prev*. 76, 159–165. doi: 10.1016/j.aap.2015.01.005

Bonnefond, M., and Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Curr. Biol.* 22, 1969–1974. doi: 10.1016/j.cub.2012.08.029

Braboszcz, C., and Delorme, A. (2011). Lost in thoughts: neural markers of low alertness during mind wandering. *Neuroimage* 54, 3040–3047. doi: 10.1016/j.neuroimage.2010.10.008

Brain Products (2018). *Brain Products GmbH / Products and Applications / actiCHamp*. Available online at: https://www.brainproducts.com/productdetails.php?id=42andtab=3 (accessed September 7, 2018).

Bureau d'Enquête et d'Analyse (2002). *Rapport Final de l'accident Survenu le 2 Octobre 1996 au Boeing 757 Exploité par Aeroperu vol Lima - Santiago*. Paris: Bureau d'Enquête et d'Analyse.

Bureau d'Enquête et d'Analyse (2012a). *Rapport Final de l'accident Survenu le 1er Juin 2009 à l'Airbus A330-203 Immatriculé F-GZCP Exploité par Air France vol AF 447 Rio de Janeiro - Paris*. Paris: Bureau d'Enquête et d'Analyse.

Bureau d'Enquête et d'Analyse (2012b). *Rapport Final sur l'incident Grave Survenu le 27 Février 2012 à l'Airbus A330 Exploité par Air France Immatriculé F-GZCG sur le vol Antananarivo - Paris*. Bureau d'Enquête et d'Analyse.

Bureau d'Enquête et d'Analyse (2016). *Rapport Final sur l'incident Grave Survenu le 24 mai 2011 sur le Dassault Falcon 7X, Immatriculé HB-JFN à Subang*. Paris: Bureau d'Enquête et d'Analyse.

Canty, A., and Ripley, B. (2017). *boot: Bootstrap R (S-Plus) Functions*. R-package, v1.3-20. Available online at: https://CRAN.R-project.org/package=boot (accessed December 16, 2020).

Carriere, J. S. A., Cheyne, J. A., and Smilek, D. (2008). Everyday attention lapses and memory failures: the affective consequences of mindlessness. *Conscious. Cogn.* 17, 835–847. doi: 10.1016/j.concog.2007.04.008

Casner, S. M., and Schooler, J. W. (2014). Thoughts in flight automation use and pilots' task-related and task-unrelated thought. *Hum. Factors* 56, 433–442. doi: 10.1177/0018720813501550

Casner, S. M., and Schooler, J. W. (2015). Vigilance impossible: diligence, distraction, and daydreaming all lead to failures in a practical monitoring task. *Conscious. Cogn.* 35, 33–41. doi: 10.1016/j.concog.2015.04.019

Cheyne, J. A., Carriere, J. S. A., and Smilek, D. (2006). Absent-mindedness: lapses of conscious awareness and everyday cognitive failures. *Conscious. Cogn.* 15, 578–592. doi: 10.1016/j.concog.2005.11.009

Cheyne, J. A., Solman, G. J. F., Carriere, J. S. A., and Smilek, D. (2009). Anatomy of an error: a bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition* 111, 98–113. doi: 10.1016/j.cognition.2008.12.009

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., and Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8719–8724. doi: 10.1073/pnas.0900234106

Clayton, M. S., Yeung, N., and Cohen Kadosh, R. (2015). The roles of cortical oscillations in sustained attention. *Trends Cogn. Sci.* 19, 188–195. doi: 10.1016/j.tics.2015.02.004

Compton, R. J., Gearinger, D., and Wild, H. (2019). The wandering mind oscillates: EEG alpha power is enhanced during moments of mind-wandering. *Cogn. Affect. Behav. Neurosci.* 19, 1184–1191. doi: 10.3758/s13415-019-00745-9

Correia, G. H., de, A., Looff, E., van Cranenburgh, S., Snelder, M., and van Arem, B. (2019). On the impact of vehicle automation on the value of travel time while performing work and leisure activities in a car: theoretical insights and results from a stated preference survey. *Transportation Res. A Policy Pract.* 119, 359–382. doi: 10.1016/j.tra.2018.11.016

Cowley, J. A. (2013). Off task thinking types and performance decrements during simulated automobile driving. *Proc. Hum. Factors Ergon. Soc. Ann. Meet.* 57, 1214–1218. doi: 10.1177/1541931213571270

Critcher, C. R., and Gilovich, T. (2010). Inferring attitudes from mind-wandering. *Pers. Soc. Psychol. Bull.* 36, 1255–1266. doi: 10.1177/0146167210375434

Crivelli, D., and Balconi, M. (2010). "Agency and inter-agency, action and joint action: theoretical and neuropsychological evidence," in *Neuropsychology of the Sense of Agency*, ed M. Balconi (Milano: Springer). doi: 10.1007/978-88-470-1587-6_6

Cummings, M. L. (2004). Automation bias in intelligent time critical decision support systems. *Am. Inst. Aeronaut. Astronaut.* 2, 557–562. doi: 10.2514/6.2004-6313

Curtis, C. E., and D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* 7, 415–423. doi: 10.1016/S1364-6613(03)00197-9

Danckert, J. (2017). Special topic introduction: understanding engagement: mind-wandering, boredom and attention. *Exp. Brain Res*. 236, 2447–2449. doi: 10.1007/s00221-017-4914-7

Dehais, F., Lafont, A., Roy, R., and Fairclough, S. (2020). A neuroergonomics approach to mental workload, engagement and human performance. *Front. Neurosci.* 14:268. doi: 10.3389/fnins.2020.00268

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Diesch, E., Andermann, M., and Rupp, A. (2012). Is the effect of tinnitus on auditory steady-state response amplitude mediated by attention? *Front. Syst. Neurosci.* 6. doi: 10.3389/fnsys.2012.00038

Dündar, C. (2015). *The effects of mind wandering on simulated driving performance* (Master's thesis). Ankara, Turkey: Middle East Technical University. Available online at: http://etd.lib.metu.edu.tr/upload/12618921/index.pdf [Accessed March 11, 2016].

Durantin, G., Dehais, F., and Delorme, A. (2015). Characterization of mind wandering using fNIRS. *Front. Syst. Neurosci.* 9:45. doi: 10.3389/fnsys.2015.00045

Eastwood, J. D., Frischen, A., Fenske, M. J., and Smilek, D. (2012). The unengaged mind: defining boredom in terms of attention. *Perspect. Psychol. Sci.* 7, 482–495. doi: 10.1177/1745691612456044

Endsley, M. R., and Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Hum. Factors* 37, 381–394. doi: 10.1518/001872095779064555

Fairclough, S. H., Gilleade, K., Ewing, K. C., and Roberts, J. (2013). Capturing user engagement via psychophysiology: measures and mechanisms for biocybernetic adaptation. *IJAACS* 6:63. doi: 10.1504/IJAACS.2013.050694

Farley, J., Risko, E. F., and Kingstone, A. (2013). Everyday attention and lecture retention: the effects of time, fidgeting, and mind wandering. *Front. Psychol.* 4:619. doi: 10.3389/fpsyg.2013.00619

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

felipeacsi, and Rietdijk, F. (2018). *Python Acoustics*. Python-package, v0.2.4. Available online at: https://github.com/python-acoustics (accessed December 16, 2020).

Foxe, J. J., Simpson, G. V., and Ahlfors, S. P. (1998). Parieto-occipital $\sim$ 10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport* 9, 3929–3933. doi: 10.1097/00001756-199812010-00030

Foxe, J. J., and Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Front. Psychol.* 2:154. doi: 10.3389/fpsyg.2011.00154

Galera, C., Orriols, L., M'Bailara, K., Laborey, M., Contrand, B., Ribereau-Gayon, R., et al. (2012). Mind wandering and driving: responsibility case-control study. *Br. Med. J.* 345, e8105–e8105. doi: 10.1136/bmj.e8105

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci.* 4, 14–21. doi: 10.1016/S1364-6613(99)01417-5

Glasauer, S., Huber, M., Basili, P., Knoll, A., and Brandt, T. (2010). "Interacting in time and space: Investigating human-human and human-robot joint action," in *19th International Symposium in Robot and Human Interactive Communication* (Viareggio, Italy: IEEE), 252–257. doi: 10.1109/ROMAN.2010.5598638

Golchert, J., Smallwood, J., Jefferies, E., Seli, P., Huntenburg, J. M., Liem, F., et al. (2016). Individual variation in intentionality in the mind-wandering state is reflected in the integration of the default-mode, fronto-parietal, and limbic networks. *Neuroimage* 146, 226–235. doi: 10.1016/j.neuroimage.2016.11.025

Gouraud, J., Delorme, A., and Berberian, B. (2018a). Influence of automation on mind wandering frequency in sustained attention. *Conscious. Cogn.* 66, 54–64. doi: 10.1016/j.concog.2018.09.012

Gouraud, J., Delorme, A., and Berberian, B. (2018b). Out of the loop, in your bubble: mind wandering is independent from automation reliability, but influences task engagement. *Front. Hum. Neurosci.* 12:383. doi: 10.3389/fnhum.2018.00383

Gramann, K., Fairclough, S. H., Zander, T. O., and Ayaz, H. (2017). Editorial: Trends in Neuroergonomics. *Front. Hum. Neurosci.* 11:165. doi: 10.3389/fnhum.2017.00165

Griskova, I., Morup, M., Parnas, J., Ruksenas, O., and Arnfred, S. M. (2007). The amplitude and phase precision of 40 Hz auditory steady-state response depend on the level of arousal. *Exp. Brain Res.* 183, 133–138. doi: 10.1007/s00221-007-1111-0

Haggard, P. (2017). Sense of agency in the human brain. *Nat. Rev. Neurosci.* 18, 196–207. doi: 10.1038/nrn.2017.14

Hancock, P. A. (2013). In search of vigilance: the problem of iatrogenically created psychological phenomena. *Am. Psychol.* 68, 97–109. doi: 10.1037/a0030214

Harb, M., Xiao, Y., Circella, G., Mokhtarian, P. L., and Walker, J. L. (2018). Projecting travelers into a world of self-driving vehicles: estimating travel behavior implications via a naturalistic experiment. *Transportation* 45, 1671–1685. doi: 10.1007/s11116-018-9937-9

He, J., Becic, E., Lee, Y.-C., and McCarley, J. S. (2011). Mind wandering behind the wheel: performance and oculomotor correlates. *Hum. Factors* 53, 17–21. doi: 10.1177/0018720810391530

Head, J., and Helton, W. S. (2016). The troubling science of neurophenomenology. *Exp. Brain Res.* 236, 2463–2467. doi: 10.1007/s00221-016-4623-7

Hosseini, A., and Lienkamp, M. (2016). "Predictive safety based on track-before-detect for teleoperated driving through communication time delay," in *2016 IEEE Intelligent Vehicles Symposium (IV)* (Gothenberg), 165–172. doi: 10.1109/IVS.2016.7535381

Jauk, E., Benedek, M., and Neubauer, A. C. (2012). Tackling creativity at its roots: evidence for different patterns of EEG alpha activity related to convergent and divergent modes of task processing. *Int. J. Psychophysiol.* 84, 219–225. doi: 10.1016/j.ijpsycho.2012.02.012

Johnson, J. A., and Zatorre, R. J. (2006). Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. *Neuroimage* 31, 1673–1681. doi: 10.1016/j.neuroimage.2006.02.026

Kam, J. W. Y. (2010). *Slow Fluctuations in Attentional Control of Sensory Cortex* (Master's thesis). Vancouver, BC: University of British Columbia.

Kam, J. W. Y., Dao, E., Blinn, P., Krigolson, O. E., Boyd, L. A., and Handy, T. C. (2012). Mind wandering and motor control: off-task thinking disrupts the online adjustment of behavior. *Front. Hum. Neurosci.* 6:329. doi: 10.3389/fnhum.2012.00329

Kam, J. W. Y., Dao, E., Farley, J., Fitzpatrick, K., Smallwood, J., Schooler, J. W., et al. (2011). Slow fluctuations in attentional control of sensory cortex. *J. Cogn. Neurosci.* 23, 460–470. doi: 10.1162/jocn.2010.21443

Kam, J. W. Y., Griffin, S., Shen, A., Patel, S., Hinrichs, H., Heinze, H.-J., et al. (2019). Systematic comparison between a wireless EEG system with dry electrodes and a wired EEG system with wet electrodes. *Neuroimage* 184, 119–129. doi: 10.1016/j.neuroimage.2018.09.012

Kam, J. W. Y., and Handy, T. C. (2013). The neurocognitive consequences of the wandering mind: a mechanistic account of sensory-motor decoupling. *Front. Psychol.* 4:725. doi: 10.3389/fpsyg.2013.00725

Kam, J. W. Y., Xu, J., and Handy, T. C. (2014). I don't feel your pain (as much): The desensitizing effect of mind wandering on the perception of others' discomfort. *Cogn. Affect. Behav. Neurosci.* 14, 286–296. doi: 10.3758/s13415-013-0197-z

Kang, L., Zhao, W., Qi, B., and Banerjee, S. (2018). "Augmenting self-driving with remote control: challenges and directions," in *Proceedings of the 19th International Workshop on Mobile Computing Systems and Applications HotMobile'18* (New York, NY: Association for Computing Machinery), 19–24. doi: 10.1145/3177102.3177104

Keitel, C., Schröger, E., Saupe, K., and Müller, M. M. (2011). Sustained selective intermodal attention modulates processing of language-like stimuli. *Exp. Brain Res.* 213, 321–327. doi: 10.1007/s00221-011-2667-2

Kelley, K. (2007a). Confidence intervals for standardized effect sizes: theory, application, and implementation. *J. Stat. Soft.* 20. doi: 10.18637/jss.v020.i08

Kelley, K. (2007b). Methods for the behavioral, educational, and social sciences: an R package. *Behav. Res. Methods* 39, 979–984. doi: 10.3758/BF03192993

Kelley, K. (2020). *The MBESS R Package*. R-package, v4.8.0. Available online at: https://www3.nd.edu/$\sim$kkelley/site/MBESS.html (accessed December 16, 2020).

Konishi, M., McLaren, D. G., Engen, H., and Smallwood, J. (2015). Shaped by the past: the default mode network supports cognition that is independent of immediate perceptual input. *PLoS ONE* 10:e0132209. doi: 10.1371/journal.pone.0132209

Krimsky, M., Forster, D. E., Llabre, M. M., and Jha, A. P. (2017). The influence of time on task on mind wandering and visual working memory. *Cognition* 169, 84–90. doi: 10.1016/j.cognition.2017.08.006

Kurihashi, S., Matsuno, Y., and Tanaka, K. (2015). "Self-perception of assister driver responsibility and contribution in mutual assistance system," in *Human Interface and the Management of Information. Information and Knowledge in Context*, ed S. Yamamoto (Cham: Springer International Publishing), 233–242. doi: 10.1007/978-3-319-20618-9_23

Lee, J. D. (2014). Dynamics of driver distraction: the process of engaging and disengaging. *Ann. Adv. Automot. Med.* 58, 24–32.

Leontiev, A. N. (2014). Activity and consciousness. *Rev. Dialectus* 2, 159–183. Available online at: http://periodicos.ufc.br/dialectus/article/view/5173

Lichiardopol, S. (2007). *A Survey on Teleoperation*. Technische Universiteit Eindhoven.

Loukopoulos, D. L. D., and Field, M. (2001). "Cockpit interruptions and distractions: a line observation study," in *Proceedings of the 11th International Symposium on Aviation Psychology* (Columbus, OH), 7.

Louw, T., Kountouriotis, G., Carsten, O., and Merat, N. (2015a). "Driver inattention during vehicle automation: how does driver engagement affect resumption of control?" in *Proceedings of the 4th International Conference on Driver Distraction and Inattention (DDI2015)* (Sydney, Australia: ARRB Group). Available online at: http://eprints.whiterose.ac.uk/91858/ (accessed August 24, 2017).

Louw, T., and Merat, N. (2017). Are you in the loop? Using gaze dispersion to understand driver visual attention during vehicle automation. *Transportation Res. C Emerg. Technol.* 76, 35–50. doi: 10.1016/j.trc.2017.01.001

Louw, T., Merat, N., and Jamson, H. (2015b). "Engaging with highly automated driving: to be or not to be in the loop," in *8th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 189–195. Available online at: https://pdfs.semanticscholar.org/3927/19a3a281e27464ed8776d75aaae82332723a.pdf (accessed August 24, 2017).

Mahajan, Y., Davis, C., and Kim, J. (2014). Attentional modulation of auditory steady-state responses. *PLoS ONE* 9:e110902. doi: 10.1371/journal.pone.0110902

Mar, R. A., Mason, M. F., and Litvack, A. (2012). How daydreaming relates to life satisfaction, loneliness, and social support: the importance of gender and daydream content. *Conscious. Cogn* 21, 401–407. doi: 10.1016/j.concog.2011.08.001

McMillan, R. L., Kaufman, S. B., and Singer, J. L. (2013). Ode to positive constructive daydreaming. *Front. Psychol.* 4:626. doi: 10.3389/fpsyg.2013.00626

McVay, J. C., and Kane, M. J. (2009). Conducting the train of thought: working memory capacity, goal neglect, and mind wandering in an executive-control task. *J. Exp. Psychol.* 35:196. doi: 10.1037/a0014104

Mullen, T. R., Kothe, C. A. E., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., et al. (2015). Real-time neuroimaging and cognitive monitoring using wearable dry EEG. *IEEE Transactions on Biomed. Eng.* 62, 2553–2567. doi: 10.1109/TBME.2015.2481482

Müller, N., Schlee, W., Hartmann, T., Lorenz, I., and Weisz, N. (2009). Top-down modulation of the auditory steady-state response in a task-switch paradigm. *Front. Hum. Neurosci.* 3:1. doi: 10.3389/neuro.09.001.2009

National Transportation Safety Board (1975). *Aircraft Accident Report, Northwest Airlines Incorporated, Boeing 727-25, N26 4US near Thiells, New York, December 1st 1974*. National Transport Safety Board.

Naujoks, F., Purucker, C., and Neukum, A. (2016). Secondary task engagement and vehicle automation – Comparing the effects of different automation levels in an on-road experiment. *Transportation Res. F Traffic Psychol. Behav.* 38, 67–82. doi: 10.1016/j.trf.2016.01.011

Neigel, A. R., Claypoole, V. L., Smith, S. L., Waldfogle, G. E., Fraulini, N. W., Hancock, G. M., et al. (2019). Engaging the human operator:

a review of the theoretical support for the vigilance decrement and a discussion of practical applications. *Theor. Issues Ergon. Sci.* 21, 1–20. doi: 10.1080/1463922X.2019.1682712

Neumeier, S., Wintersberger, P., Frison, A.-K., Becher, A., Facchi, C., and Riener, A. (2019). "Teleoperation: the holy grail to solve problems of automated driving? sure, but latency matters," in *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Utrecht Netherlands: ACM), 186–197. doi: 10.1145/3342197.3344534

Nielsen, C. W., Goodrich, M. A., and Ricks, R. W. (2007). Ecological interfaces for improving mobile robot teleoperation. *IEEE Trans. Robot.* 23, 927–941. doi: 10.1109/TRO.2007.907479

Norman, S. D., and Orlady, H. W. (1988). "Flight deck automation: promises and realities," *Proceedings of a NASA/FAA/Industry Workshop*. Available online at: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19900004068.pdf (accessed April 21, 2020).

Obhi, S. S., and Hall, P. (2011). Sense of agency in joint action: influence of human and computer co-actors. *Exp. Brain Res.* 211, 663–670. doi: 10.1007/s00221-011-2662-7

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., and Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20 s before they occur. *J. Neurosci.* 29, 8604–8611. doi: 10.1523/JNEUROSCI.5967-08.2009

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2010). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011:156869. doi: 10.1155/2011/156869

Oostenveld, R., and Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clin. Neurophysiol.* 112, 713–719. doi: 10.1016/S1388-2457(00)00527-7

OpenBCI (2016). OpenBCI - open source biosensing tools (EEG, EMG, EKG, and more). Available online at: http://openbci.com/ (accessed May 23, 2016).

Ottaviani, C., and Couyoumdjian, A. (2013). Pros and cons of a wandering mind: a prospective study. *Front. Psychol.* 4:524. doi: 10.3389/fpsyg.2013.00524

Parasuraman, R. (2011). Neuroergonomics: brain, cognition, and performance at work. *Curr. Dir. Psychol. Sci.* 20, 181–186. doi: 10.1177/0963721411409176

Pattyn, N., Neyt, X., Henderickx, D., and Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiol. Behav.* 93, 369–378. doi: 10.1016/j.physbeh.2007.09.016

Picton, T. W., John, M. S., Dimitrijevic, A., and Purcell, D. (2003). Human auditory steady-state responses. *Int. J. Audiol.* 42, 177–219. doi: 10.3109/14992020309101316

Porcu, E., Keitel, C., and Müller, M. M. (2014). Visual, auditory and tactile stimuli compete for early sensory processing capacities within but not between senses. *Neuroimage* 97, 224–235. doi: 10.1016/j.neuroimage.2014.04.024

Psychology Software Tools (2018). *E-Prime* ®. Sharpsburg, PA, v2.0. Available online at: https://pstnet.com/products/e-prime/ (accessed September 8, 2018).

Python Software Foundation (2018). *Python., v3.6*. Available online at: https://www.python.org/ (accessed December 16, 2020).

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/ (accessed December 16, 2020).

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682. doi: 10.1073/pnas.98.2.676

Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., and Kingstone, A. (2012). Everyday attention: variation in mind wandering and memory in a lecture: mind wandering. *Appl. Cogn. Psychol.* 26, 234–242. doi: 10.1002/acp.1814

RStudio Team (2015). *RStudio: Integrated Development for R*. Boston, MA: RStudio Inc., v1.0.143. Available online at: http://www.rstudio.com/ (accessed December 16, 2020).

SAE International (2018). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-road Motor Vehicles*. Society of Automotive Engineers International.

Sahaï, A., Pacherie, E., Grynszpan, O., and Berberian, B. (2017). Predictive mechanisms are not involved the same way during human-human vs. human-machine interactions: a review. *Front. Neurorobot.* 11:52. doi: 10.3389/fnbot.2017.00052

Salomone, M., Burle, B., Fabre, L., and Berberian, B. (2021). An electromyographic analysis of the effects of cognitive fatigue on online and anticipatory action control. *Front. Hum. Neurosci.* 14:615046. doi: 10.3389/fnhum.2020.615046

Saupe, K., Schröger, E., Andersen, S., and Müller, M. M. (2009a). Neural mechanisms of intermodal sustained selective attention with concurrently presented auditory and visual stimuli. *Front. Hum. Neurosci.* 3:58. doi: 10.3389/neuro.09.058.2009

Saupe, K., Widmann, A., Bendixen, A., Müller, M. M., and Schröger, E. (2009b). Effects of intermodal attention on the auditory steady-state response and the event-related potential. *Psychophysiology* 46, 321–327. doi: 10.1111/j.1469-8986.2008.00765.x

Schad, D. J., Nuthmann, A., and Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: objective measures reveal mindless reading at different levels. *Cognition* 125, 179–194. doi: 10.1016/j.cognition.2012.07.004

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., and Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends Cogn. Sci.* 15, 319–326. doi: 10.1016/j.tics.2011.05.006

Seli, P., Kane, M., Smallwood, J., Schacter, D. L., Maillet, D., Schooler, J., et al. (2018). Mind-wandering as a natural kind: a family-resemblances view. *Trends Cogn. Sci.* 22, 479–490. doi: 10.1016/j.tics.2018.03.010

Singer, J. L. (1975). *The Inner World of Daydreaming*. Harper & Row.

Skosnik, P. D., Krishnan, G. P., and O'Donnell, B. F. (2007). The effect of selective attention on the gamma-band auditory steady-state response. *Neurosci. Lett.* 420, 223–228. doi: 10.1016/j.neulet.2007.04.072

Smallwood, J. (2011). Mind-wandering while reading: attentional decoupling, mindless reading and the cascade model of inattention. *Lang. Linguis. Compass* 5, 63–77. doi: 10.1111/j.1749-818X.2010.00263.x

Smallwood, J., Baracaia, S. F., Lowe, M., and Obonsawin, M. (2003). Task unrelated thought whilst encoding information. *Conscious. Cogn.* 12, 452–484. doi: 10.1016/S1053-8100(03)00018-7

Smallwood, J., McSpadden, M., and Schooler, J. W. (2008). When attention matters: the curious incident of the wandering mind. *Mem. Cogn.* 36, 1144–1150. doi: 10.3758/MC.36.6.1144

Smallwood, J., Obonsawin, M., and Reid, H. (2002). The effects of block duration and task demands on the experience of task unrelated thought. *Imagin. Cogn. Pers.* 22, 13–31. doi: 10.2190/TBML-N8JN-W5YB-4L9R

Smallwood, J., O'Connor, R. C., Sudberry, M. V., Haskell, C., and Ballantyne, C. (2004). The consequences of encoding information on the maintenance of internally generated images and thoughts: the role of meaning complexes. *Conscious. Cogn.* 13, 789–820. doi: 10.1016/j.concog.2004.07.004

Smallwood, J., Riby, L., Heim, D., and Davies, J. B. (2006). Encoding during the attentional lapse: accuracy of encoding during the semantic sustained attention to response task. *Conscious. Cogn.* 15, 218–231. doi: 10.1016/j.concog.2005.03.003

Smallwood, J., and Schooler, J. W. (2006). The restless mind. *Psychol. Bull.* 132, 946–958. doi: 10.1037/0033-2909.132.6.946

Smallwood, J., and Schooler, J. W. (2015). The science of mind wandering: empirically navigating the stream of consciousness. *Annu. Rev. Psychol.* 66, 487–518. doi: 10.1146/annurev-psych-010814-015331

SmartCap (2020). *SmartCap Technologies | Measure Alertness. Eliminate Fatigue*. SmartCapTech. Available at: http://www.smartcaptech.com/ (accessed September 27, 2020).

Stawarczyk, D., Majerus, S., Maquet, P., and D'Argembeau, A. (2011). Neural correlates of ongoing conscious experience: both task-unrelatedness and stimulus-independence are related to default network activity. *PLoS ONE* 6:e16997. doi: 10.1371/journal.pone.0016997

Terhune, D. B., Croucher, M., Marcusson-Clavertz, D., and Macdonald, J. S. P. (2017). Time contracts and temporal precision declines when the mind wanders. *J. Exp. Psychol.* 43, 1864–1871. doi: 10.1037/xhp0000461

The Mathworks Inc. (1992). *MATLAB User's Guide*. Natick, MA: The Mathworks Inc.

This Place (2016). MindRDR. Available online at: http://mindrdr.thisplace.com/static/index.html (accessed May 23, 2016).

Thomson, D. R., Seli, P., Besner, D., and Smilek, D. (2014). On the link between mind wandering and task performance over time. *Conscious. Cogn.* 27, 14–26. doi: 10.1016/j.concog.2014.04.001

Uanhoro, J. O. (2017). *Effect Size Calculators*. Available online at: https://effect-size-calculator.herokuapp.com/ (accessed August 30, 2018).

van der Wal, C. N., and Irrmischer, M. (2015). "Myndplay: measuring attention regulation with single dry electrode brain computer interface," in *Proceedings of the 8th International Conference on Brain Informatics and Health* (London: Springer, Cham), 192–201. doi: 10.1007/978-3-319-23344-4_19

Villena-González, M., López, V., and Rodríguez, E. (2016). Orienting attention to visual or verbal/auditory imagery differentially impairs the processing of visual stimuli. *Neuroimage* 132, 71–78. doi: 10.1016/j.neuroimage.2016.02.013

Voicikas, A., Niciute, I., Ruksenas, O., and Griskova-Bulanova, I. (2016). Effect of attention on 40 Hz auditory steady-state response depends on the stimulation type: flutter amplitude modulated tones versus clicks. *Neurosci. Lett.* 629, 215–220. doi: 10.1016/j.neulet.2016.07.019

Wen, W., and Haggard, P. (2018). Control changes the way we look at the world. *J. Cogn. Neurosci.* 30, 603–619. doi: 10.1162/jocn_a_01226

Westgate, E. C., and Wilson, T. D. (2018). Boring thoughts and bored minds: the MAC model of boredom and cognitive engagement. *Psychol. Rev.* 125, 689–713. doi: 10.1037/rev0000097

Wohlschläger, A., Engbert, K., and Haggard, P. (2003a). Intentionality as a constituting condition for the own self—and other selves. *Conscious. Cogn.* 12, 708–716. doi: 10.1016/S1053-8100(03)00083-7

Wohlschläger, A., Haggard, P., Gesierich, B., and Prinz, W. (2003b). The perceived onset time of self-and other-generated actions. *Psychol. Sci.* 14, 586–591. doi: 10.1046/j.0956-7976.2003.psci_1469.x

Yao, D., Wang, L., Oostenveld, R., Nielsen, K. D., Arendt-Nielsen, L., and Chen, A. C. N. (2005). A comparative study of different references for EEG spectral mapping: the issue of the neutral reference and the use of the infinity reference. *Physiol. Meas.* 26, 173–184. doi: 10.1088/0967-3334/26/3/003

Zedelius, C. M., and Schooler, J. W. (2016). The richness of inner experience: relating styles of daydreaming to creative processes. *Front. Psychol.* 6:2063. doi: 10.3389/fpsyg.2015.02063

Zhang, T. (2020). Toward automated vehicle teleoperation: vision, opportunities, and challenges. *IEEE Internet Things J.* 7, 11347–11354. doi: 10.1109/JIOT.2020.3028766

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroergonomics

ORIGINAL RESEARCH
published: 13 August 2021
doi: 10.3389/fnrgo.2021.691993

Neuroplastic Reorganization Induced by Sensory Augmentation for Self-Localization During Locomotion

Hiroyuki Sakai [1]*, *Sayako Ueda* [2]*, *Kenichi Ueno* [3] *and Takatsune Kumada* [4]

[1] Human Science Laboratory, Toyota Central R&D Laboratories, Inc., Tokyo, Japan, [2] TOYOTA Collaboration Center, RIKEN Center for Brain Science, Wako, Japan, [3] Support Unit for Functional Magnetic Resonance Imaging, RIKEN Center for Brain Science, Wako, Japan, [4] Graduate School of Informatics, Kyoto University, Kyoto, Japan

Sensory skills can be augmented through training and technological support. This process is underpinned by neural plasticity in the brain. We previously demonstrated that auditory-based sensory augmentation can be used to assist self-localization during locomotion. However, the neural mechanisms underlying this phenomenon remain unclear. Here, by using functional magnetic resonance imaging, we aimed to identify the neuroplastic reorganization induced by sensory augmentation training for self-localization during locomotion. We compared activation in response to auditory cues for self-localization before, the day after, and 1 month after 8 days of sensory augmentation training in a simulated driving environment. Self-localization accuracy improved after sensory augmentation training, compared with the control (normal driving) condition; importantly, sensory augmentation training resulted in auditory responses not only in temporal auditory areas but also in higher-order somatosensory areas extending to the supramarginal gyrus and the parietal operculum. This sensory reorganization had disappeared by 1 month after the end of the training. These results suggest that the use of auditory cues for self-localization during locomotion relies on multimodality in higher-order somatosensory areas, despite substantial evidence that information for self-localization during driving is estimated from visual cues on the proximal part of the road. Our findings imply that the involvement of higher-order somatosensory, rather than visual, areas is crucial for acquiring augmented sensory skills for self-localization during locomotion.

Keywords: augmentation, plasticity, driving, fMRI, locomotion

OPEN ACCESS

Edited by:
Lewis L. Chuang,
Ludwig Maximilian University of Munich, Germany

Reviewed by:
Emanuelle Reynaud,
Université de Lyon, France
Giacinto Barresi,
Italian Institute of Technology (IIT), Italy

***Correspondence:**
Hiroyuki Sakai
sakai@mosk.tytlabs.co.jp
Sayako Ueda
sayako.ueda@riken.jp

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: *07 April 2021*
Accepted: *21 July 2021*
Published: *13 August 2021*

Citation:
Sakai H, Ueda S, Ueno K and Kumada T (2021) Neuroplastic Reorganization Induced by Sensory Augmentation for Self-Localization During Locomotion. Front. Neuroergon. 2:691993. doi: 10.3389/fnrgo.2021.691993

1. INTRODUCTION

1.1. Background

Sensory skills can be augmented through training and technological support. An obvious example is Braille reading, in which well-trained individuals can read letters via tactile sensation when touching Braille symbols. Various devices have been developed to facilitate sensory augmentation. Such devices detect environmental information by using electronic sensors and convert it into stimuli delivered to a sensory organ that is not innately associated with the information. For example, one device translates visual scenes recorded by a digital camera into auditory stimuli by converting elevation to pitch and brightness to loudness (Meijer, 1992). After training with such a device, users are able to discriminate several visual objects without actually seeing them (Striem-Amit et al., 2012). Another research group has developed a waist-belt-type vibration device that constantly displays magnetic north (Nagel et al., 2005). This enables users to utilize a newly acquired magnetic orientation sense to navigate in outdoor environments.

Sensory augmentation is underpinned by neuroplastic reorganization in the brain. A pioneering study by Sadato et al. (1996) demonstrated the activation of visual cortical areas in blind individuals during Braille reading. Similar cortical reorganization in visual cortical areas has been observed after training of blind individuals in distance perception aided by an ultrasound echolocation device (De Volder et al., 1999) or in letter recognition by using an electrotactile stimulation device (Ptito et al., 2005). Such reorganization has also been observed after training of sighted individuals in depth perception by using a device that converts visual scenes into auditory stimuli (Renier et al., 2005), and of both sighted and blind individuals in object recognition with a visual-to-auditory sensory substitution device (Amedi et al., 2007). In the case of this last device, it has been further demonstrated that shape information conveyed by auditory stimuli activates different visual areas, depending on the category of the object (Reich et al., 2011; Striem-Amit et al., 2012; Abboud et al., 2015).

Recently, we proposed a novel sensory augmentation system that assists self-localization during vehicle driving (Ueda et al., 2019). This system translates a vehicle's lateral position in a traffic lane into binaural balance of white-noise loudness, enabling drivers to sense a lane line they are approaching as increased loudness to the ipsilateral ear and decreased loudness to the contralateral ear. By using this auditory-based self-localization assistance system for locomotion, we demonstrated in a simulated driving environment that drivers developed the ability to control the vehicle accurately by using auditory cues, even when the visual information needed to estimate vehicle lateral position (i.e., the proximal part of the road) was occluded. This was the first successful attempt to show the applicability of sensory augmentation to time-sensitive daily-life situations in healthy individuals. However, the underlying neural mechanisms of the training effects remain unclear.

1.2. Objective

Here, by using functional magnetic resonance imaging (fMRI), we aimed to identify the neuroplastic reorganization induced by sensory augmentation training for self-localization during locomotion. Specifically, we employed a pretest-training-posttest design comprising three separate fMRI sessions (pretest, posttest, and follow-up test), before and after driver training with or without the sensory augmentation assistance system in a simulated environment. In the pretest fMRI session, activation of the response to auditory stimuli conveying vehicle lateral position information was investigated by using a conventional block design protocol. After the pretest fMRI session, participants were randomly assigned to one of two training conditions (normal driving [ND] or sensory augmentation [SA] conditions) and accordingly performed 8 days of driver training. On the day after the last training day, a posttest fMRI session was conducted by using a protocol identical to that applied in the pretest. The fMRI session was further repeated approximately 1 month later as a follow-up test. We then identified training-related changes in auditory responses under each condition.

Previous studies examining neuroplasticity after sensory augmentation have reported a clear tendency for trained sensory information to become processed in the brain areas for a sensory modality associated with the content of the information, rather than the carrier of the information. Therefore, we expected that vehicle lateral position translated into auditory stimuli would be processed in visual areas after training, because vehicle lateral position is considered to be estimated from the visual cues contained in driving scenes as viewed from the driver's perspective (Land and Horwood, 1995; Billington et al., 2010; Frissen and Mars, 2014).

2. MATERIALS AND METHODS

2.1. Participants

Fourteen adults (2 females; 12 males) aged 20–28 years participated in the study and received financial compensation for their participation. All participants self-reported normal or corrected-to-normal vision and normal hearing. In addition, all participants were free from psychiatric, neurological, and major medical illnesses, as determined by medical history. They were all right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). Each participant provided written informed consent. Experimental protocols were approved by the RIKEN Research Ethics Committee [Wako3 28-17(4)] and were conducted according to the principles of the Declaration of Helsinki.

2.2. Study Design

To identify neuroplastic reorganization associated with sensory augmentation for self-localization during locomotion, we employed a pretest-training-posttest design comprising three separate fMRI sessions (pretest, posttest, and follow-up). After the pretest fMRI scan, participants were randomly assigned to either the ND or SA condition, and they accordingly performed eight sequential days of training (excluding weekends and occasional absences) in a simulated environment. On the day after the last training day, a posttest fMRI scan was conducted. In addition, another follow-up fMRI scan was performed ∼1 month after the end of the training. To minimize observer effects, the experimenters conducting the fMRI sessions were blind to which participant was assigned to which training condition.

2.3. Driving Training

For sensory augmentation training, we used a custom-made driving simulator comprising a fixed-base cockpit (GTD-SPECi, Rossomodello Co., Ltd., Tomioka, Japan), a force-feedback steering device (T500RS, Guillemot Corp., Carentoir, France), and a 60-inch LCD monitor (LC60XL10, Sharp Corp., Sakai, Japan) located in front of the cockpit (**Figure 1A**). The system was identical to the one used in our previous behavioral study (Ueda et al., 2019). Driving scenes as viewed from the driver's perspective were reconstructed at 60 Hz and displayed on the

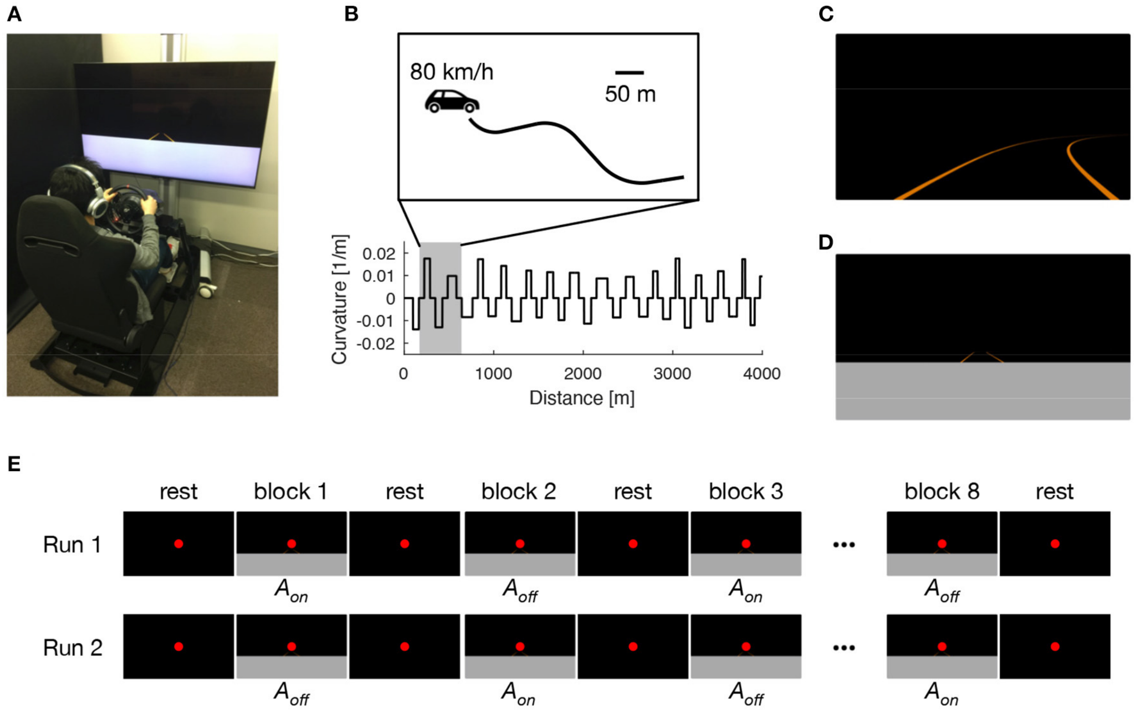


FIGURE 1 | Experimental setup. Participants performed 8 days of lane-keeping training on a custom-made driving simulator **(A)**. Driving courses consisted of a winding road that alternately curved leftward and rightward with a random curvature between 1/120 and 1/60 m^{-1}, interleaved with straight sections with a constant length of 60 m **(B)**. Under normal driving (ND) condition, driving scenes were presented with auditory white noise with a constant loudness to both ears via headphones **(C)**. Under sensory augmentation (SA) condition, the lower part of the driving scene was occluded to restrict the visual cues needed for estimating vehicle lateral position, and information regarding vehicle lateral position was instead provided via binaural auditory stimuli **(D)**. The pretest and posttest MRI sessions comprised two block-design runs, in which participants were exposed to sensory stimuli that were relevant to the lane-keeping task under SA condition. In each run, driving scenes were presented alternately with (A_{on}) and without (A_{off}) auditory cues for vehicle lateral position **(E)**.

monitor, subtending horizontal and vertical visual angles of 66° and 43°, respectively, at a viewing distance of 105 cm.

On each of the 8 training days, participants performed 20 trials of a lane-keeping task (180 s for each trial) under their assigned training condition (i.e., ND or SA condition). In each trial, participants were required to keep their vehicle in the center of a traffic lane by using the steering wheel. The traffic lane was defined by left and right lane lines, giving a lane width of 3.5 m. The lane alternately curved leftward and rightward with a random curvature between 1/120 and 1/60 m^{-1}; this was interleaved with straight sections with a constant length of 60 m (**Figure 1B**). The vehicle automatically traveled with a constant speed of 80 km/h and, therefore, no pedal operations were needed. No other road users were present throughout the task. Under ND condition (**Figure 1C**), white noise with a constant loudness (−50 dB attenuated from the maximum level that we predetermined to be comfortably tolerable to participants) was presented to both ears of participants via headphones during the entire lane-keeping task. In contrast, under SA condition (**Figure 1D**), the lower part of the driving scenes (i.e., the proximal part of the road) was occluded to restrict the visual cues needed to estimate vehicle lateral position. Instead, information regarding vehicle lateral position was provided via binaural auditory stimuli. Specifically, leftward (rightward) deviation of vehicle lateral position from the center of the lane was signaled by increasing the loudness of white noise in the left (right) ear and decreasing the loudness of white noise in the right (left) ear with a constant gain of 25 dB/m; in the center of the lane, the loudness level was equal in both ears (−50 dB). Participants assigned to SA condition were instructed in advance regarding the meaning of the auditory stimuli. Under both conditions, the entire experiment, including preparation and rest breaks (< 1 min) between trials, took <90 min on each training day.

Driving performance in the lane-keeping task was evaluated in terms of accuracy and smoothness of vehicle control by using the standard deviation of vehicle lateral position (SDLP) and the maximum steering wheel velocity during curve negotiation (SWV), respectively. There is accumulating evidence that SDLP reflects compensatory steering control that makes use of the visual information provided by the proximal part of the road, whereas SWV reflects anticipatory steering control that makes use of the information provided by the more distant part of the road (Frissen and Mars, 2014; Ueda et al., 2019). For both SDLP and SWV, lower values represent better performance. To examine the effects of the training on driving performance, we estimated the learning plateau and learning rate under each training condition. First, we computed the trajectory of each performance metric (i.e., SDLP or SWV) as a function of trial number (total, 160 trials) for each participant and fitted an exponential function [$Y = a + b\exp(-cX)$ where Y is a performance metric, X is the trial number, and a, b, and c are regression parameters] to the mean trajectory averaged across participants to estimate the learning plateau (a) and learning rate (c) with a 95% confidence interval (CI).

2.4. MRI Data Acquisition

MRI data acquisition was performed on a Siemens 3T Prisma scanner with a 64-channel head array coil (Siemens Medical System, Erlangen, Germany). In each fMRI session, a high-resolution T1-weighted structural image was acquired by using a 3D MPRAGE sequence (Mugler and Brookeman, 1990) with an echo time (TE) of 3.25 ms, a repetition time (TR) of 1,700 ms, an inversion time (TI) of 0.9 s, a flip angle (FA) of 8°, a field of view (FOV) of 256 × 256 mm^2, a matrix size of 256 × 256, a slice thickness of 1 mm, 192 contiguous sagittal slices, and an acceleration factor of 2 for the GRAPPA parallel imaging technique (Griswold et al., 2002). Then, fMRI data were collected by using a gradient echo T2*-weighted echo-planar imaging (EPI) sequence with a TE of 30 ms, a TR of 1,556 ms, an FA of 74°, an FOV of 200 × 200 mm^2, a matrix size of 100 × 100, a slice thickness of 2 mm, 72 contiguous axial slices, an acceleration factor of 2 for the GRAPPA parallel imaging technique, and a multi-band factor of 3.

During fMRI, participants were exposed to sensory stimuli that were relevant to the lane-keeping task under SA condition (**Figure 1E**). Specifically, after a 16-s rest period, participants were given a 16-s stimulation period followed by a 16-s rest period 8 times (giving a "run" of 272 s in total, corresponding to 175 EPI volumes). For the stimulation periods, driving video clips were created from driving log data in which an experimenter (HS) had performed lane-keeping under SA training condition. The video clips were displayed on a translucent screen with visual angles of 26.3 and 15.6°. In half of the video clips in each run, the audio tracks containing auditory cues for vehicle lateral position were present (A_{on}). In the other half, the audio tracks were removed (A_{off}). These two kinds of video clips were presented alternately in a run. Throughout the run, regardless of whether the participant was in a rest period or a stimulation period, a red dot was presented constantly for fixation. The run was presented twice to each participant, with the first run starting with A_{on} after the first rest period and the second starting with A_{off} instead. In each run, participants were asked to answer which lane line (left or right) was being approached by pressing either the left or right button of a hand-held switch box with their right index finger or middle finger, respectively. The purpose of this instruction was to maintain the participant's attention on the sensory stimuli, rather than to evaluate performance in the scanner. Throughout the entire fMRI session, we confirmed that participants were in a state of arousal by monitoring their eyes via a camera. In addition, their respiratory and cardiac signals were collected by using a pressure sensor and a pulse oximeter, respectively. These electrophysiological signals were used afterwards to remove physiological fluctuations from the EPI images (Hu et al., 1995).

2.5. fMRI Data Analysis

Task fMRI data were preprocessed for each participant using SPM12 (v7487; www.fil.ion.ucl.ac.uk/spm) with the CAT12 toolbox (r1184; www.neuro.uni-jena.de/cat). First, to obtain a deformation field for accurate normalization of EPI images, a T1-weighted structural image collected in the pretest session was processed by using a segmentation procedure in CAT12. Second, EPI images were preprocessed by using SPM12. All EPI images acquired in the three fMRI sessions were realigned to the first image in the pretest session for head motion correction and then coregistered to the structural image in the pretest session. The coregistered EPI images were normalized to MNI (Montreal Neurological Institute) space by using the deformation field and then smoothed with an isotropic Gaussian kernel with a full width at half maximum of 6 mm.

To identify neuroplastic reorganization induced by the sensory augmentation training, we performed a voxelwise general linear model analysis. At the first level, we modeled the two stimulus effects (A_{on} and A_{off}) for each run by using boxcar functions convolved with the canonical hemodynamic response function. By contrasting the two stimulus conditions ($A_{on} > A_{off}$) for each fMRI session, we obtained activation in response to auditory cues for vehicle lateral position. At the second level, we entered these individual contrast images into a 2-by-3 full factorial model with a between-subjects factor of training condition (ND, SA) and a within-subjects factor of fMRI session (pretest, posttest, follow-up). We then performed a conjunction analysis of auditory responses in all fMRI sessions (pretest ∩ posttest ∩ follow-up) to identify the brain regions consistently activated before and after the lane-keeping training. We also compared auditory responses between different fMRI sessions to identify brain regions that were more active after training, compared with before training (i.e., pretest < posttest and pretest < follow-up). For all analyses, the results were considered statistically significant at $P < 0.01$ cluster-level family-wise error (FWE) corrected for multiple comparisons, with a voxel-level threshold of $P < 0.005$ uncorrected.

3. RESULTS

3.1. Driving Training

Under both ND and SA conditions, lane-keeping accuracy as assessed by using SDLP improved as training progressed (**Figure 2A**). Regression analysis revealed that learning rate for lane-keeping accuracy was slower under SA condition (0.036; 95% CI, 0.29–0.42) than under ND condition (0.055; 95% CI, 0.044–0.067), and that the learning plateau was lower under SA condition (0.31; 95% CI, 0.30–0.32) than under ND condition (0.37; 95% CI, 0.37–0.38). Lane-keeping smoothness, as assessed by using SWV, also improved as training progressed, under both conditions (**Figure 2B**). Regression analysis revealed that learning rate for lane-keeping smoothness was slower under SA condition (0.20; 95% CI, 0.15–0.25) than under ND condition (0.55; 95% CI, 0.32–0.78) and that the learning plateaus were comparable under SA (0.057; 95% CI, 0.056–0.057) and ND (0.056; 95% CI, 0.055–0.056) conditions. Overall, lane-keeping performance under SA condition improved more slowly during

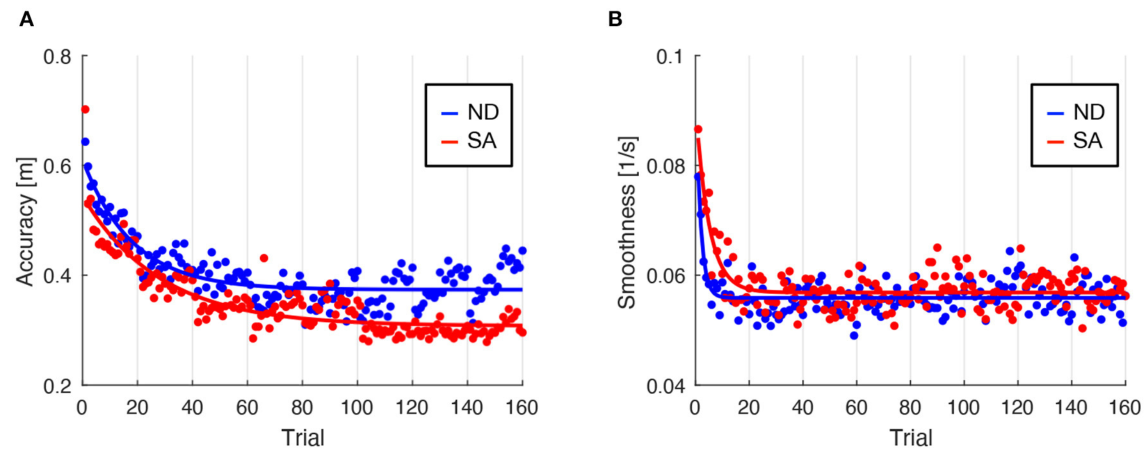


FIGURE 2 | Trajectories of driving performance during lane-keeping training. Driving performance was indexed with regard to the accuracy **(A)** and smoothness **(B)** of steering as a function of trial number. Blue and red dots represent mean performance in each trial, separately averaged across participants under ND and SA conditions. Each line denotes a learning curve fitted to determine learning rate and learning plateau. Vertical thin lines indicate respective training days.

training, but after training it was eventually comparable to, or better than, that under ND condition, even though for SA participants the visual information essential for lane-keeping was unavailable.

3.2. Brain Activation

During fMRI, participants kept their eyes open except for natural blinking. In addition, all participants pressed the buttons more than 3 times in every task block (**Supplementary Figure 1**). These results suggest that the participants were engaged in the task.

Conjunction analysis of auditory responses across the three fMRI sessions (i.e., pretest, posttest, and follow-up) revealed significant brain activation in the superior temporal gyri bilaterally under both ND and SA conditions (**Table 1**), although the left clusters did not satisfy the statistical criteria under ND condition. These results indicate that auditory cues for vehicle lateral position consistently activated temporal auditory areas.

Training-induced changes (pretest < posttest) in auditory responses differed between SA and ND conditions (**Table 1**). Under ND condition, comparison of pretest and posttest auditory responses revealed increased activation in the pre-supplementary motor area (pre-SMA) and anterior insular cortex (**Figure 3A**). In contrast, under SA condition, increased activation in the posttest fMRI compared with the pretest fMRI was found in the somatosensory areas bilaterally, including in the parietal operculum, supramarginal gyrus (SMG), and postcentral gyrus (**Figure 4A**), which were adjacent to, but not overlapping with, the superior temporal auditory areas revealed by the conjunction analysis. Under both conditions, no significant clusters were found in the follow-up fMRI compared with the pretest fMRI. These changes in the posttest fMRI results can be regarded as the signatures of neuroplastic reorganization induced by the lane-keeping training.

Furthermore, we examined the relationships between training-induced neuroplastic changes (pretest < posttest) and driving performance achieved by the training (i.e., the learning plateau in lane-keeping accuracy; **Figure 2A**). In the somatosensory areas, increased activation after the training was negatively correlated with lane-keeping accuracy under SA condition (left cluster, $r = -0.69$, $P = 0.044$; right cluster, $r = -0.36$, $P = 0.21$; one-tailed t-test) but positively correlated under ND condition (left cluster, $r = 0.51$, $P = 0.12$; right cluster, $r = 0.79$, $P = 0.017$; **Figures 4B,C**), suggesting that better lane-keeping performance under SA condition was associated with greater involvement of the somatosensory areas in auditory processing after the training. A similar tendency was also found in both the pre-SMA (ND, $r = 0.86$, $P = 0.0069$; SA, $r = -0.35$, $P = 0.22$; **Figure 3B**) and the anterior insular cortex (ND, $r = 0.78$, $P = 0.019$; SA, $r = -0.59$, $P = 0.083$; **Figure 3C**), suggesting that better lane-keeping performance under ND condition was associated with less involvement of these frontal areas in auditory processing after the training. In all clusters, between-group differences in correlation coefficients were statistically significant (for the left somatosensory cluster, $P = 0.023$; for the right somatosensory cluster, $P = 0.020$; for the pre-SMA cluster, $P = 0.010$; for the anterior insular cluster, $P = 0.0073$; one-tailed Z-test).

4. DISCUSSION

Increasing attention has been paid to sensory augmentation for not only sensory impaired but also healthy individuals, because it can open new horizons for human-machine/computer interface development by reconsidering human-environment interactions (Di Pino et al., 2014). We previously demonstrated the potential of an auditory-based self-localization assistance system for locomotion in a simulated driving environment; this demonstration pioneered the application of sensory augmentation to time-sensitive daily-life situations in healthy individuals (Ueda et al., 2019). However, the neural mechanisms

TABLE 1 | Brain regions activated in response to auditory cues for self-localization.

Region	x	y	z	t	Size
(ND, pretest ∩ posttest ∩ follow-up)					
Superior temporal gyrus	50	−28	12	5.95	511
Superior temporal gyrus	−34	−36	16	4.05	111†
Superior temporal gyrus	−56	−36	12	3.49	79†
(SA, pretest ∩ posttest ∩ follow-up)					
Superior temporal gyrus	66	−32	12	7.09	824
Superior temporal gyrus	−62	−38	10	6.48	806
(ND, pretest < posttest)					
Pre-supplementary motor area	10	12	56	5.76	1284
Anterior insular cortex	−36	8	6	5.63	392
(SA, pretest < posttest)					
Parietal operculum	−60	−28	18	3.91	372
Supramarginal gyrus	−58	−28	28	3.73	
Postcentral gyrus	−56	−28	52	3.60	
Postcentral gyrus	64	−16	22	3.76	392
Supramarginal gyrus	66	−24	24	3.62	

Statistical significance was set at cluster-level P < 0.01, family-wise error (FWE) corrected for multiple comparisons; the clusters marked with a dagger (†) were not large enough to satisfy this criterion. ND and SA represent normal driving and sensory augmentation training conditions, respectively. Cluster locations are given in Montreal Neurological Institute coordinates.

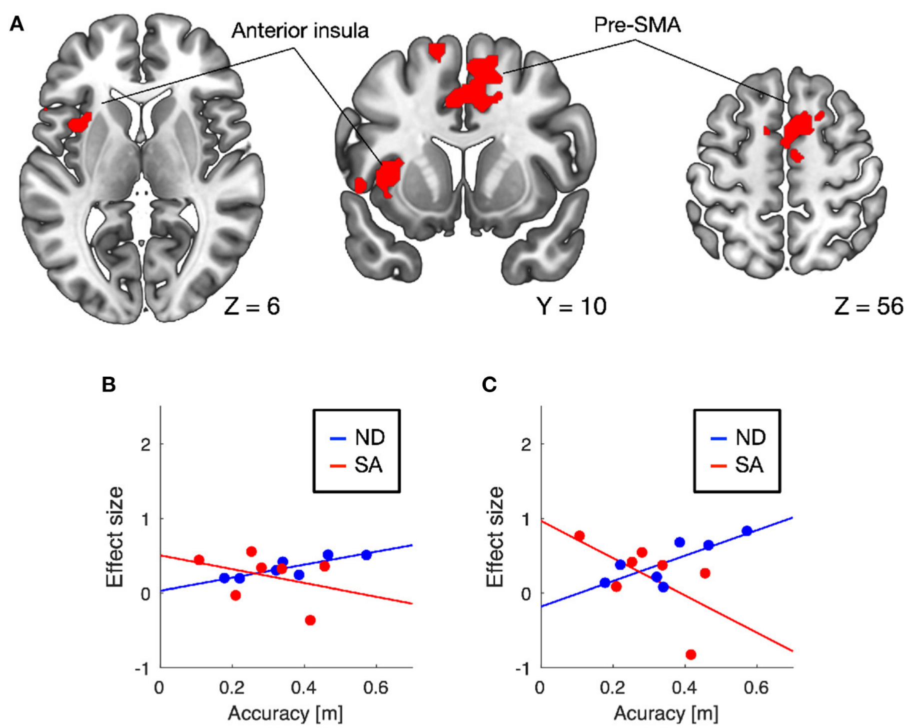


FIGURE 3 | Brain regions showing increased auditory responses after lane-keeping training under ND condition (pretest < posttest). Significant clusters (family-wise error (FWE) corrected $P < 0.01$) were found in the pre-supplementary motor area (pre-SMA) and left anterior insula **(A)**. In both clusters, increased activation after the training was positively correlated with lane-keeping accuracy under ND condition, but negatively under SA condition (**B**, pre-SMA; **C**, anterior insula).

underlying sensory augmentation for self-localization during locomotion remain unclear. Identifying the neural underpinnings of augmentation could provide information that would improve our self-localization assistance systems and facilitate the development of novel sensory augmentation devices to assist locomotion.

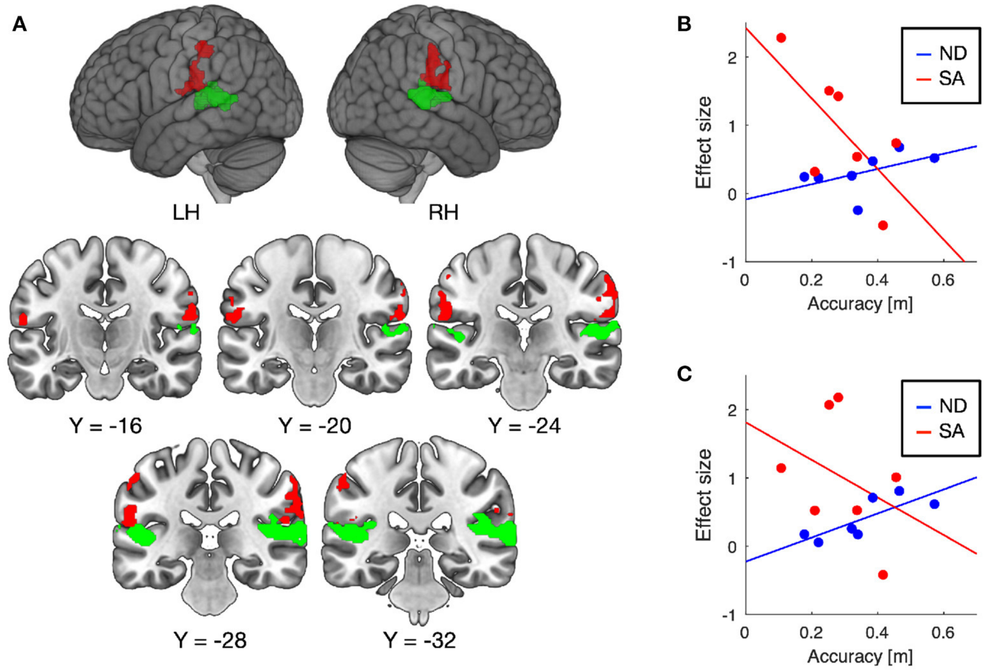


FIGURE 4 | Brain regions showing increased auditory responses after lane-keeping training under SA condition (pretest < posttest) **(A)**. Significant clusters (FWE-corrected $P < 0.01$) were found in the somatosensory areas bilaterally (red), which were not overlapped with the superior temporal auditory areas (green) revealed by the conjunction analysis (i.e., pretest ∩ posttest ∩ follow-up). In both clusters, increased activation after the training was negatively correlated with lane-keeping accuracy under SA condition, but positively under ND condition (**B**, left cluster; **C**, right cluster).

Here, we measured brain activation in response to auditory cues on vehicle lateral position before and after lane-keeping training to identify training-induced neuroplastic reorganization. Although we expected auditory responses in the occipital visual areas after the training, we found training-induced increases in the activation of somatosensory areas in the parietal lobe. We also found that greater involvement of somatosensory areas in auditory processing after the training was associated with better lane-keeping performance under SA condition. Not only the parietal operculum, which is considered to be the location of the secondary somatosensory cortex (Ruben et al., 2001; Eickhoff et al., 2006), but also the SMG is involved in somatosensory processing as a human homologue of the tertiary somatosensory cortex in the monkey (Caselli, 1993; Hagen and Pardo, 2002). Our findings appear to be consistent with the neuroplastic changes involved in auditory-induced somatosensory sensation after stroke. Beauchamp and Ro (2008) investigated auditory responses in a stroke patient complaining of sound-touch synesthesia and found substantial activation within the secondary somatosensory cortex in response to auditory stimuli. Furthermore, there is growing evidence of the inherent multimodal nature of the higher-order somatosensory cortices. Bremmer et al. (2001) demonstrated that the SMG was activated consistently by visual, auditory, and tactile motion stimuli. More recently, Pérez-Bellido et al. (2018) reported that auditory frequency information was represented in somatosensory areas, including the SMG, as well as in the auditory cortices. These findings clearly indicate that auditory information can be processed in higher-order somatosensory cortices. Taken together, our data suggest that the use of auditory cues for self-localization during locomotion relies on multimodality in higher-order somatosensory cortices rather than the occipital visual cortices, even though the vehicle lateral position conveyed by the auditory cues in this study is usually estimated from visual cues within the proximal part of the road (Land and Horwood, 1995; Frissen and Mars, 2014; Ueda et al., 2019).

However, another interpretation is possible because there were differences in visual as well as auditory stimuli during driving training between ND and SA conditions. Specifically, the lower part of the driving scenes (i.e., the proximal part of the road) was visually occluded in SA, but not ND, condition. Therefore, the results could be interpreted as indicating that improved driving performance in SA condition resulted from the completion of occluded visual information rather than the complementary use of auditory cues (sensory augmentation). Although this interpretation cannot be ruled out in the current experiment, our previous behavioral study (Ueda et al., 2019) revealed that when the proximal part of the road is occluded, driving performance does not improve without auditory cues for self-localization. Thus, we concluded that it is more likely that somatosensory involvement in auditory processing after SA training resulted from the acquisition of augmented sensory skill using auditory cues for self-localization rather than visual completion.

It could be argued that this sensory augmentation for self-localization during locomotion resembles echolocation and therefore that they are likely to share neural substrates. Echolocation is an augmented sensory skill by which the sound reverberation of mouth clicks is used to infer spatial information in the surrounding environment, and it can be used for navigation by blind people. From a functional point of view, echolocation and the sensory augmentation that we examined here can both be regarded as auditory-based localization skills. However, their neural bases appear to differ from each other. Several lines of evidence demonstrate the involvement of occipital visual, rather than somatosensory, areas in echolocation (Thaler et al., 2011, 2014; Wallmeier et al., 2015). Furthermore, Thaler and Foresteire (2017) reported that, in sighted individuals, echolocation performance was disrupted by task-unrelated visual, but not tactile, stimuli, also suggesting that there is a lack of involvement of the somatosensory areas in echolocation. The major difference between our findings and those of these previous echolocation studies is in the action demanded for locomotion. Our participants were required to control a vehicle by using augmented sensory information for self-localization, whereas in the previous studies participants were asked to extract spatial information on the surrounding objects from echoes. The use of augmented sensory information for self-localization during locomotion might be crucial for the involvement of somatosensory areas. In fact, one previous study reported the echo-related activation of extensive parietal areas in a situation where the use of echolocation ability to detect path directions during walking was required (Fiehler et al., 2015). The neuroplastic reorganization induced by sensory augmentation may depend not only on what kinds of content are conveyed via an augmented sense, but also on how the augmented sensory information is used to accomplish task demands. This notion requires further clarification in future research.

Another important finding of our study was that superior lane-keeping accuracy was achieved with, rather than without, the auditory-based self-localization assistance system. It is evident that the visual information contained in the proximal part of the road is typically critical to accurate lane-keeping (Land and Horwood, 1995; Frissen and Mars, 2014). By using the same experimental setup, we previously demonstrated that when the proximal part of the road was occluded, lane-keeping accuracy was markedly degraded and did not improve with training (Ueda et al., 2019). However, we observed that the learning curve of lane-keeping accuracy reached a better plateau under SA condition than under ND condition. This result may be attributable to more accurate and precise feedback of lane-keeping errors under SA condition. In general, sensorimotor learning tasks require sensory error signals if a person is to achieve fine motor control (Kawato et al., 1987). Under ND condition and in typical visual-based vehicle driving, drivers are required to estimate the vehicle lateral position by using visual cues from the proximal part of the road. In contrast, the auditory cues used under SA condition can provide exact information about the vehicle lateral position. Nevertheless, in terms of both accuracy and smoothness performance metrics, learning was slower under SA than under ND condition. This presumably reflects the additional cognitive cost of utilizing auditory cues for self-localization.

We also found training-induced neuroplastic changes under ND condition. This was an unexpected result, because under ND condition the auditory stimuli presented during lane-keeping training were totally irrelevant to the task. A possible interpretation for this is that the cognitive control required to ignore the task-irrelevant auditory stimuli during lane-keeping training resulted in training-induced greater activation of the pre-SMA and the anterior insular cortex. In fact, both these regions are considered part of the cognitive control network (Cole and Schneider, 2007; Niendam et al., 2012). In addition, Smucny et al. (2013) showed that the pre-SMA is engaged when auditory distraction is present during a highly demanding cognitive task. The anterior insular cortex is also known to play a crucial role in suppressing distractor inference (Bunge et al., 2002). This interpretation seems to dovetail with lesser involvement of the superior temporal auditory cortex under ND condition than under SA condition (**Table 1**). Under this interpretation, furthermore, more involvement of the abovementioned frontal regions under ND condition would indicate the assignment of more cognitive resources to distractor inference suppression; this is consistent with our finding that under ND condition, lane-keeping performance was negatively associated with increased involvement of these areas after the training.

Several limitations of our study should be noted. First, the small sample size restricts the ability to generalize our findings. In particular, we identified training-induced neuroplastic changes only as within-group differences, not as between-group differences (i.e., interaction between training condition and fMRI session). Replication with a larger sample size is needed to improve the generalizability of our results. Second, brain activation associated with sensory augmentation for self-localization needs to be investigated by using a more realistic driving environment. We made a substantial effort to expose participants in the MRI scanner to sensory stimuli similar to those experienced during sensory augmentation training. However, the sensory stimulation remained different between inside and outside the scanner (e.g., the size of the visual stimuli). In addition, we examined brain activation during passive exposure to sensory stimuli, but not during active behaviors using augmented sensory information. According to previous neuroimaging studies (Uchiyama et al., 2003, 2012), active compared with passive driving additionally activates not only motor areas but also various sensory and association areas. It will be important for future studies to investigate the role of somatosensory responses to auditory cues for self-localization in such driving-related networks. Because there is empirical evidence that brain activation associated with sensory augmentation is context-dependent (Sadato et al., 1996), differences in stimuli and behavior might have influenced our findings.

5. CONCLUSION

We demonstrated here that sensory augmentation for self-localization during locomotion results in extensive neuroplastic reorganization in the human cerebral cortex. Interestingly, although we expected training-induced reorganization in the occipital visual areas, we observed neuroplastic changes in higher-order somatosensory areas. This finding suggests that the involvement of somatosensory, rather than visual, areas is crucial for acquiring augmented sensory skills for self-localization during locomotion, even though self-localization is considered to rely heavily on vision. Our data also showed that, depending on how it is used, sensory augmentation can enable better performance in healthy individuals, particularly in situations where the information provided by an augmented sense initially requires the complex computation of sensory signals. Our findings will facilitate further applications of sensory augmentation to human-computer/machine interface development.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by RIKEN Research Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

HS, SU, and TK contributed to conception and design of the study. SU and KU performed the experiment. HS and SU performed the data analysis and wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

ACKNOWLEDGMENTS

We thank Drs. Chisato Suzuki and Kenji Haruhana from RIKEN Center for Brain Science for helping MRI data acquisition.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnrgo.2021.691993/full#supplementary-material

REFERENCES

Abboud, S., Maidenbaum, S., Dehaene, S., and Amedi, A. (2015). A number-form area in the blind. *Nat. Commun.* 6:6026. doi: 10.1038/ncomms7026

Amedi, A., Stern, W. M., Camprodon, J. A., Bermpohl, F., Merabet, L., Rotman, S., et al. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nat. Neurosci.* 10, 687–689. doi: 10.1038/nn1912

Beauchamp, M. S., and Ro, T. (2008). Neural substrates of sound-touch synesthesia after a thalamic lesion. *J. Neurosci.* 28, 13696–13702. doi: 10.1523/JNEUROSCI.3872-08.2008

Billington, J., Field, D. T., Wilkie, R. M., and Wann, J. P. (2010). An fMRI study of parietal cortex involvement in the visual guidance of locomotion. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1495–1507. doi: 10.1037/a0018728

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K.-P., et al. (2001). Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron* 29, 287–296. doi: 10.1016/S0896-6273(01)00198-2

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., and Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron* 33, 301–311. doi: 10.1016/S0896-6273(01)00583-9

Caselli, R. J. (1993). Ventrolateral and dorsomedial somatosensory association cortex damage produces distinct somesthetic syndromes in humans. *Neurology* 43, 762–771. doi: 10.1212/WNL.43.4.762

Cole, M. W., and Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *Neuroimage* 37, 343–360. doi: 10.1016/j.neuroimage.2007.03.071

De Volder, A. G., Catalan-Ahumada, M., Robert, A., Bol, A., Labar, D., Coppens, A., et al. (1999). Changes in occipital cortex activity in early blind humans using a sensory substitution device. *Brain Res.* 826, 128–134. doi: 10.1016/S0006-8993(99)01275-5

Di Pino, G., Maravita, A., Zollo, L., Guglielmelli, E., and Di Lazzaro, V. (2014). Augmentation-related brain plasticity. *Front. Sys. Neurosci.* 8:109. doi: 10.3389/fnsys.2014.00109

Eickhoff, S. B., Amunts, K., Mohlberg, H., and Zilles, K. (2006). The human parietal operculum. II. Stereotaxic maps and correlation with functional imaging results. *Cereb. Cortex* 16, 268–279. doi: 10.1093/cercor/bhi106

Fiehler, K., Schütz, I., Meller, T., and Thaler, L. (2015). Neural correlates of human echolocation of path direction during walking. *Multisens. Res.* 28, 195–226. doi: 10.1163/22134808-00002491

Frissen, I., and Mars, F. (2014). The effect of visual degradation on anticipatory and compensatory steering control. *Q. J. Exp. Psychol.* 67, 499–507. doi: 10.1080/17470218.2013.819518

Griswold, M. A., Jakob, P. M., Heidemann, R. M., Nittka, M., Jellus, V., Wang, J., et al. (2002). Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn. Reson. Med.* 47, 1202–1210. doi: 10.1002/mrm.10171

Hagen, M. C., and Pardo, J. V. (2002). PET studies of somatosensory processing of light touch. *Behav. Brain Res.* 135, 133–140. doi: 10.1016/S0166-4328(02)00142-0

Hu, X., Le, T. H., Parrish, T., and Erhard, P. (1995). Retrospective estimation and correction of physiological fluctuation in functional MRI. *Magn. Reson. Med.* 34, 201–212. doi: 10.1002/mrm.1910340211

Kawato, M., Furukawa, K., and Suzuki, R. (1987). A hierarchical neural-network model for control and learning of voluntary movement. *Biol. Cybern.* 57, 169–185. doi: 10.1007/BF00364149

Land, M., and Horwood, J. (1995). Which parts of the road guide steering? *Nature* 377, 339–340. doi: 10.1038/377339a0

Meijer, P. B. (1992). An experimental system for auditory image representations. *IEEE Trans. Biomed. Eng.* 39, 112–121. doi: 10.1109/10.121642

Mugler, J. P., and Brookeman, J. R. (1990). Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). *Magn. Reson. Med.* 15, 152–157. doi: 10.1002/mrm.1910150117

Nagel, S. K., Carl, C., Kringe, T., Märtin, R., and König, P. (2005). Beyond sensory substitution-learning the sixth sense. *J. Neural Eng.* 2, R13–R26. doi: 10.1088/1741-2560/2/4/R02

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., and Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn. Affect. Behav. Neurosci.* 12, 241–268. doi: 10.3758/s13415-011-0083-5

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Pérez-Bellido, A., Anne Barnes, K., Crommett, L. E., and Yau, J. M. (2018). Auditory frequency representations in human somatosensory cortex. *Cereb. Cortex* 28, 3908–3921. doi: 10.1093/cercor/bhx255

Ptito, M., Moesgaard, S. M., Gjedde, A., and Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain* 128(Pt 3), 606–614. doi: 10.1093/brain/awh380

Reich, L., Szwed, M., Cohen, L., and Amedi, A. (2011). A ventral visual stream reading center independent of visual experience. *Curr. Biol.* 21, 363–368. doi: 10.1016/j.cub.2011.01.040

Renier, L., Collignon, O., Poirier, C., Tranduy, D., Vanlierde, A., Bol, A., et al. (2005). Cross-modal activation of visual cortex during depth perception using auditory substitution of vision. *Neuroimage* 26, 573–580. doi: 10.1016/j.neuroimage.2005.01.047

Ruben, J., Schwiemann, J., Deuchert, M., Meyer, R., Krause, T., Curio, G., et al. (2001). Somatotopic organization of human secondary somatosensory cortex. *Cereb. Cortex* 11, 463–473. doi: 10.1093/cercor/11.5.463

Sadato, N., Pascual-Leone, A., Grafman, J., Ibañez, V., Deiber, M. P., Dold, G., et al. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature* 380, 526–528. doi: 10.1038/380526a0

Smucny, J., Rojas, D. C., Eichman, L. C., and Tregellas, J. R. (2013). Neuronal effects of auditory distraction on visual attention. *Brain Cogn.* 81, 263–270. doi: 10.1016/j.bandc.2012.11.008

Striem-Amit, E., Cohen, L., Dehaene, S., and Amedi, A. (2012). Reading with sounds: sensory substitution selectively activates the visual word form area in the blind. *Neuron* 76, 640–652. doi: 10.1016/j.neuron.2012.08.026

Thaler, L., Arnott, S. R., and Goodale, M. A. (2011). Neural correlates of natural human echolocation in early and late blind echolocation experts. *PLos ONE* 6:e20162. doi: 10.1371/journal.pone.0020162

Thaler, L., and Foresteire, D. (2017). Visual sensory stimulation interferes with people's ability to echolocate object size. *Sci. Rep.* 7:13069. doi: 10.1038/s41598-017-12967-3

Thaler, L., Milne, J. L., Arnott, S. R., Kish, D., and Goodale, M. A. (2014). Neural correlates of motion processing through echolocation, source hearing, and vision in blind echolocation experts and sighted echolocation novices. *J. Neurophysiol.* 111, 112–127. doi: 10.1152/jn.00501.2013

Uchiyama, Y., Ebe, K., Kozato, A., Okada, T., and Sadato, N. (2003). The neural substrates of driving at a safe distance: a functional MRI study. *Neurosci. Lett.* 352, 199–202. doi: 10.1016/j.neulet.2003.08.072

Uchiyama, Y., Toyoda, H., Sakai, H., Shin, D., Ebe, K., and Sadato, N. (2012). Suppression of brain activity related to a car-following task with an auditory task: an fMRI study. *Trans. Res. Part F* 15, 25–37. doi: 10.1016/j.trf.2011.11.002

Ueda, S., Sakai, H., and Kumada, T. (2019). A novel approach to sensorimotor skill acquisition utilizing sensory substitution: a driving simulation study. *Sci. Rep.* 9:17886. doi: 10.1038/s41598-019-54324-6

Wallmeier, L., Kish, D., Wiegrebe, L., and Flanagin, V. L. (2015). Aural localization of silent objects by active human biosonar: neural representations of virtual echo-acoustic space. *Eur. J. Neurosci.* 41, 533–545. doi: 10.1111/ejn.12843

Conflict of Interest: The authors declare that this study received funding from Toyota Motor Corporation. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication. In addition, HS was employed by Toyota Central R&D Laboratories.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroergonomics

ORIGINAL RESEARCH
published: 14 September 2021
doi: 10.3389/fnrgo.2021.729197

Measuring Correlates of Mental Workload During Simulated Driving Using cEEGrid Electrodes: A Test–Retest Reliability Analysis

Stephan Getzmann[1], Julian E. Reiser[1], Melanie Karthaus[1], Georg Rudinger[2] and Edmund Wascher[1]*

[1] *IfADo - Leibniz Research Centre for Working Environment and Human Factors, Dortmund, Germany,* [2] *Uzbonn - Society for Empirical Social Research and Evaluation, Bonn, Germany*

OPEN ACCESS

Edited by:
Christer Ahlström,
Swedish National Road and Transport Research Institute, Sweden

Reviewed by:
Gianluca Di Flumeri,
Sapienza University of Rome, Italy
Yu-Kai Wang,
University of Technology Sydney, Australia
Amy McDonnell,
The University of Utah, United States

***Correspondence:**
Stephan Getzmann
getzmann@ifado.de

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: *22 June 2021*
Accepted: *17 August 2021*
Published: *14 September 2021*

Citation:
Getzmann S, Reiser JE, Karthaus M, Rudinger G and Wascher E (2021) Measuring Correlates of Mental Workload During Simulated Driving Using cEEGrid Electrodes: A Test–Retest Reliability Analysis. Front. Neuroergon. 2:729197. doi: 10.3389/fnrgo.2021.729197

The EEG reflects mental processes, especially modulations in the alpha and theta frequency bands are associated with attention and the allocation of mental resources. EEG has also been used to study mental processes while driving, both in real environments and in virtual reality. However, conventional EEG methods are of limited use outside of controlled laboratory settings. While modern EEG technologies offer hardly any restrictions for the user, they often still have limitations in measurement reliability. We recently showed that low-density EEG methods using film-based round the ear electrodes (cEEGrids) are well-suited to map mental processes while driving a car in a driving simulator. In the present follow-up study, we explored aspects of ecological and internal validity of the cEEGrid measurements. We analyzed longitudinal data of 127 adults, who drove the same driving course in a virtual environment twice at intervals of 12–15 months while the EEG was recorded. Modulations in the alpha and theta frequency bands as well as within behavioral parameters (driving speed and steering wheel angular velocity) which were highly consistent over the two measurement time points were found to reflect the complexity of the driving task. At the intraindividual level, small to moderate (albeit significant) correlations were observed in about 2/3 of the participants, while other participants showed significant deviations between the two measurements. Thus, the test-retest reliability at the intra-individual level was rather low and challenges the value of the application for diagnostic purposes. However, across all participants the reliability and ecological validity of cEEGrid electrodes were satisfactory in the context of driving-related parameters.

Keywords: EEG, driving, mental work load, cEEGrids, test-retest reliability

INTRODUCTION

Neurophysiological research methods have a long tradition of deriving mental processes both under laboratory conditions and in real-life environments. While in the first case a high degree of experimental control and reliability of measurements is can be assumed, measurements of neurophysiological parameters in the field (still) represent a challenge, but also an opportunity toward a higher ecological validity (Engel et al., 2013; Parada, 2018; Parada and Rossi, 2020). Especially with regard to EEG, the development of modern recording methods and analysis routines has opened up completely new possibilities to map the work of the brain under real conditions (for a recent review, Wascher et al., 2021). In two recent studies, for example, we showed that mental workload during the processing of cognitive tasks while walking on differently challenging courses was not only reflected in performance measures, but that it was also associated with modulations in brain activity (Reiser et al., 2019, 2020). While both studies clearly demonstrated the usability of EEG measurements under out-of-laboratory everyday conditions, conventional electrode caps have been used here, which offer a good prerequisite for EEG recording, but are unfavorable in real-life environments for many reasons: they are conspicuous, time-consuming to apply, restrict the user's mobility, and are of limited use when high ecological validity is important – especially when possible influence of the measurement method on the measurement results should be minimized (e.g., Sterr et al., 2018; Mikkelsen et al., 2019).

Alternative solutions are provided by new recording technologies. The use of dry electrodes, for example, is such a technology, which has proven to be very reliable, but easier to apply and wear compared to conventional wet electrodes (Di Flumeri et al., 2019). Even more inconspicuous is the cEEGrid system, in which the EEG is recorded by only a few film-based round the ear electrodes. The cEEGrids technology not only avoids restrictions arising from conventional electrode setups (Symeonidou et al., 2018), but is also easier and faster to apply than conventional multichannel Cap-EEG. At the same time, they offer a sufficient signal quality and allow for valid and reliable measurements (Mirkovic et al., 2016; Bleichner and Debener, 2017). Previous research has shown, for example, that it is possible to derive neurophysiological correlates of cognitive processes from the oscillatory brain activity recorded via cEEGrid electrodes both in an auditory oddball task (Debener et al., 2015) and a visual Simon task (Pacharra et al., 2017).

The good practicability of the cEEGrids technology was only recently demonstrated in a large-scale study on driving abilities of seniors, in which older adults drove an ∼1-h close-to-reality driving simulator course, consisting of different road sections with various challenges for the driver (Wascher et al., 2019). Using behavioral (driving speed, steering wheel angular velocity) and neurophysiological measures (EEG oscillatory power in the theta and alpha band frequencies), it was possible to estimate mental workload while driving, based only on characteristics of the driving situation. They found that with increasing track difficulty the steering angular velocity increased while driving speed decreased. A similar pattern was found on the electrophysiological level, whereas relative theta power increased and relative alpha power decreased. Finally, using a track-frequency analysis, it was possible to map modulations in EEG spectral power to the difficulty of the traffic situation, which highly corresponded with a priori expert ratings. This highlights the connection of behavioral and electrophysiological measures, as the findings are in line with the assumption that, firstly, reduced alpha power is a correlate of increased mental workload (Wascher et al., 2016) and attentional engagement (Pattyn et al., 2008), and, secondly, increased theta power is related to mental processing demands (Lal and Craig, 2001; Borghini et al., 2014) and associated with higher workload (Wilson and Hankins, 1994; Gevins et al., 1997) or task engagement (Yamada, 1998; Onton et al., 2005). However, the cognitive processes represented by alpha and theta activity cannot be considered separately. Especially in natural environments, for example, when driving a car (Di Flumeri et al., 2018) and when multi-tasking is required (Puma et al., 2018), numerous subtasks have to be performed, which are represented differently in oscillatory brain activity. It has been proposed that visual processing, information-gathering, and early attention allocation seems to be represented more by alpha activity, while higher cognitive processes such as integration of information, problem solving, and executive functions seem to be represented more by theta activity (Berka et al., 2007). This is also reflected in the topography, with alpha activity typically derived over parietal and theta activity over fronto-central areas (e.g., Wang et al., 2018; for review, Klimesch, 1999). Accordingly, by combining driving parameters and oscillatory activity in the alpha and theta frequency bands derived over parietal and frontal areas, respectively, it has recently been demonstrated that the current workload of a driver can reliably be determined using a mobile EEG system (Islam et al., 2020). Taken together, both measures demonstrated the flexible allocation of cognitive resources depending on the route section and difficulty (Borghini et al., 2014; Karthaus et al., 2018; for review, Lohani et al., 2019).

Results like these are overall promising, but lead toward a still unanswered question: to what extent are these EEG measurements reliable? This arising question of EEG test-retest reliability is nothing new, as studies on resting-state EEG proved that the normal EEG can be treated as an intraindividually rather stable trait (e.g., Gasser et al., 1985; Van Albada et al., 2007; Angelidis et al., 2016), with test-retest reliabilities in healthy adults typically exceeding 0.80 over intervals of more than 1 year (Hatz et al., 2015). Adding to this, task-related EEG which maps changes in cognitive states related to, for example, task difficulty was also found to have a high test-retest reliability. An exemplary study was conducted by McEvoy et al. (2000), in which subjects performed cognitive tasks at intervals of 7 days, resulting in high intraindividual correlations in oscillatory brain activity in the theta and alpha frequency bands. Comparably high reliabilities were also found in other works (e.g., Fernández et al., 1993; Fallgatter et al., 2002; Näpflin et al., 2008). In the context of driving, a study on the reproducibility of EEG modulations as consequence of driver fatigue showed high test-retest reliability as well (Lal and Craig, 2005). However, transient fluctuations in mental states like alertness and vigilance are hard to control especially under less structured experimental conditions and have typically been associated with reduced test-retest reliabilities – a pattern typically found in natural environments (Fernández et al., 1993). For mobile EEG systems, only few findings are available so far. A study in which the test-retest reliability of a single-channel, wireless EEG system was tested in healthy individuals showed reduced, but still satisfactory reliabilities over short (1-day) and longer (1-week and 1-month) retest-intervals, with Intra-Class Correlations for a group of older adults ranging between 0.51 and 0.89 in an eyes-open condition (Rogers et al., 2016). A study on cEEGrids demonstrated a sufficient test-retest reliability when measuring resting-state and task-related EEG in an auditory oddball paradigm over many hours (Debener et al., 2015).

The aim of the present study was to evaluate the test-retest reliability of the cEEGrid technology under less favorable recording conditions over an even longer time interval. For this purpose, the data of the first measurement point of our driving study presented in Wascher et al. (2019) were compared with those of the second measurement more than 1 year later. All data analyzed here were taken from a (still ongoing) large-scale investigation of the driving abilities of older adults aged between 67 and 76 years, which is designed as a longitudinal study with the same individuals being tested several times at intervals of 12–15 months. In addition to several neuropsychological tests, the project also comprises a simulated driving test during which the EEG is recorded using cEEGrid technology. The comparison between the two time points of measurement was performed on the behavioral (i.e., driving speed and steering wheel angular velocity) and EEG data (alpha and theta power) as well as their dependencies on the characteristics of the driving route. In addition, it was assessed to what extent interindividual differences could be replicated regarding the allocation of mental resources as a function of workload. Thus, while our former study demonstrated that task-related modulations of driving behavior and EEG—previously found in controlled lab settings—are also observable in a naturalistic driving simulation and cEEGrids measurements, now we focused on the following questions: (1) How have the performance parameters assessed during the driving course (i.e., driving speed and steering wheel angular velocity) changed compared to the first measurement point? (2) Can the previously found dependence of relative alpha and theta power on track difficulty be replicated at a between-subject level? (3) How strong is the intraindividual correspondence of the oscillatory measures in dependence on the track difficulty?

METHODS

Participants

All participants were part of a large-scale longitudinal investigation of the driving abilities of older adults which started in 2016. One hundred twenty-seven participants took also part in both measurement time points, completed the required driving distance twice and provided a sufficient data quality in the EEG (see below). These 127 participants (mean age 72.2 years, age range 68–77 years; 22.0 % female) all had a valid driving license and reported to be experienced drivers with an average annual mileage between 5,000 and 10,000 km/year. They had normal or corrected to normal vision and reported an overall good health status. They completed a battery of neuropsychological tests which will not be reported here. Before starting the experiment, all participants provided written informed consent. The study was approved by the local ethics committee of the Leibniz Research Centre for Working Environment and Human Factors.

Task and Procedure

The task and the experimental procedure were exactly the same for measurement points 1 and 2 (MP1 and MP2). Between MP1 and MP2, there was an average of 398.17 days (minimum 350, maximum 580, SD 37.79; about 13 months). The test procedure and data analysis have been described in detail in Wascher et al. (2019). In brief: After completing various questionnaires and performing a battery of neuropsychological followed by a vision test, the participants completed a pre-test drive lasting about 15 min. The driving route of the pre-test drive was not part of the actual test drive and intended to familiarize the drivers with the characteristics of the vehicle, its steering and braking behavior, and the static driving simulator (ST Sim, St Software B.V. Groningen, NL). Then the cEEGrid electrodes were attached and the participants completed a driving course which resembled a regular German driving test consisting of four different road sections: a section of state road with several intersections, roundabouts, and a foggy passage (SR1) was followed by a longer freeway section including several roadwork sites and a freeway parking area had to be passed (FW). This was followed by another section of state road with several left and right turn intersections (SR2), before the drivers entered the city where traffic lights, pedestrians, and cyclists had to be attended to (CT). Acoustic (verbal) and visual navigation information guided the drivers through the ∼37-km driving course. Given that not all participants finished the complete course, only the first 30 km were analyzed here.

In order to test how the mental workload of the driver was modulated by the characteristics of the driving route, the driving scenario was a priori subdivided into three driving profiles, being either simple (undisturbed ride on a free route), complex (junctions with turning, roundabouts, left turns, traffic lights, motorway entrances and exits), or interactive (interactions with other traffic participants, like overtaking or driving behind a vehicle ahead). In total, sections of simple, complex, and interactive driving profiles comprised ∼13, 8, and 9 km, respectively. These driving profiles were classified by an expert according to their assumed mental demands as of low, medium, and high task load (cf. Pauzié, 2008; Rahman et al., 2017). It should be noted, however, that this subdivision was done across all road sections (i.e., state road, freeway, and city sections), since the proportion of different route profiles was distributed rather unevenly across the road sections. Since the driving distance had to be limited to a reasonable level (also in view of the background of the study and the age of the participants), the data basis did not appear to be sufficient for a more fine-grained differentiation.

Data Recording and Processing

EEG was recorded using cEEGrids, consisting of flex-printed, C-shaped electrode arrays with 10 silver printed electrodes (Debener et al., 2015; Bleichner et al., 2016; Mirkovic et al., 2016; Pacharra et al., 2017). The cEEGrids are positioned around the participant's left and right ear using an adhesive surface (**Figure 1**). In contrast to conventional electrode setups, cEEGrids are barely visible, comfortable to wear, require only a small amount of electrode gel, and are therefore fast and

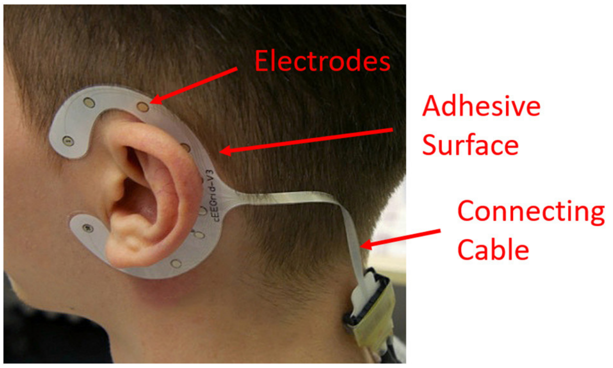


FIGURE 1 | cEEGrids technology, consisting of C-shaped electrode arrays with 10 electrodes placed around the participant's left and right ears (photo: IfADo).

easy to apply and remove. The cEEGrids were connected to a QuickAmp DC-amplifier with an on-line low-pass filtering at 280 Hz. Data were sampled at 1 kHz with a resolution of 24 bits. The two electrodes in the middle of the right cEEGrid served as ground and online reference respectively (R4a, R4b). EEG data were stored together with the driving simulator data from which driving speed and steering wheel angular velocity were derived offline. Driving speed was defined as the distance (in meters) traveled per time (in seconds) over a distance of 10 m and converted in kilometers per hour (km/h). Steering wheel angular velocity was defined as the angular speed at which the drivers turned the steering wheel, averaged over a distance of 10 m and converted in degrees per second (deg/s). In general, steering wheel angular velocity is considered an indicator of task load while driving (e.g., Antin et al., 1990; Verwey and Veltman, 1996).

The EEG analysis procedure is described in detail in Wascher et al. (2019) and is therefore only outlined briefly here. Firstly, data were checked for integrity, so that data sets with either incomplete driving distance or corrupt transmission of simulator data into the EEG recording files were discarded. After resampling to 200 Hz and band-pass filtering (1–40 Hz) of the EEG and simulator data, single EEG channels were checked for integrity by using the EEGLAB implemented rej_channel function (normed data; criterion: 4 standard deviations) to detect and discard faulty channels. Only datasets with intact reference channels after channel rejection were kept for further analyses. They were re-referenced to the average of L4b and R4b and entered into the artifact subspace reconstruction (ASR) procedure (Mullen et al., 2014, 2017). ASR is a component-based method and was proven in a number of studies (e.g., Plechawska-Wojcik et al., 2019) including a driving simulator study (Chang et al., 2019) to be highly effective in automatic filtering transient or large-amplitude artifacts (like produced by eye blinks and eye movements) from EEG data. Followingly, a time frequency decomposition was performed on each channel by convolving the data with complex Morlet wavelets. Spectral power estimates were calculated as the squared absolute values of the complex convolution result and were averaged across channels. Finally, participants with total EEG power that deviated by more than 3 standard deviations from the median were discarded and the complete data set was excluded. In total, the 127 participants (described in section Participants) who had complete data sets at both measurement points were included into the further analysis.

Data Analysis

We conducted two different approaches to assess the retest reliability of the EEG data, first a *task-load related analysis*, investigating whether the EEG measures at both measurement points depended on the driving profile in the same way, and second an (intra-individual) *correlational analysis*, comparing the EEG measures along the route at MP1 and MP2 separately for each subject. In addition to the spectral power in theta (3–6 Hz) and alpha (7–10 Hz) frequency bands, behavioral data (driving speed and steering wheel angle velocity) were analyzed to test whether behavioral results reflect the same pattern as the EEG results. It should be noted that we chose a lower than typical frequency range for determining alpha activity. The reason for this is the shift in alpha activity toward lower frequencies that is often observed with increasing age (e.g., Van Albada et al., 2010; Chiang et al., 2011). In our earlier analysis, we also measured a mean alpha frequency of <9 Hz and therefore chose the frequency range of 7–10 Hz (Wascher et al., 2019). For this reason, and also for reasons of better comparability with our previous study, we have maintained this frequency range here as well.

In the task-load related analysis, behavioral and EEG data were averaged across the driving course, separately for simple, complex, and interactive driving profiles, and mean values were entered into 2 × 3 ANOVAs with measurement point (MP1, MP2) and driving profile (simple, complex, interactive) as within-subjects factors. Effect size estimates (adj η_p^2) are reported according to Mordkoff (2019). As in our former study, not only raw power values of alpha and theta activity were analyzed, but also relative power values, representing the percentage of the power in a given frequency band relative to the total power. We therefore calculated the contribution of each frequency to the overall signal by applying a vector normalization across all frequencies for each time point. The result were the so-called alpha and theta *fractions*. The idea behind this normalization is that high power and high variance in oscillatory activity across all frequency ranges often masks effects in the alpha and theta regions, which may become more prominent by forming the relative power values. Thus, there is evidence that relative power fluctuations are more related to experimental effects than absolute power fluctuations (Klimesch, 1999; Kilner et al., 2005; Labounek et al., 2015).

For the correlational analysis, we conducted the track-frequency analysis (as detailed in Wascher et al., 2019), in which the time period of the EEG recording was mapped onto the 30-km driving route using 43 predefined landmarks for each participant. The landmarks consisted of defined route points to which a trigger was written into the EEG recording as soon as the vehicle passed this point. For the sections between the landmarks, the waypoints were estimated from the current speed of the vehicle at that point. Thus, we achieved a temporal-spatial assignment, in which each time point of the EEG measurement was assigned to a track section by stretching and compressing the EEG data in the temporal domain. For the track-frequency analysis, 3,000 10-meter track segments were generated, covering the entire 30-km driving route. To determine the alpha and theta power along the track, first a time-frequency analysis was performed over all time points. Based on this analysis, the mean power fraction for alpha and theta power was calculated for each of the 10-meter track segments, then z-transformed across all the 3,000 data points and low-pass filtered by a $\pm$ 40 m moving average. The 95% confidence intervals were calculated and are shown in **Figure 4** for MP1 and MP2. In order to determine the relationships of oscillatory power measured at the two measurement points on an intraindividual level, correlations between MP1 and MP2 were computed across the entire 30-km driving course. That is, Pearson's r correlations between the alpha and theta power values measured at MP1 and MP2 were computed across the 3,000 10-meter track segments for each participant. Associations between the two measurement points were regarded as weak, moderate, or high for correlation coefficients of 0.10, 0.30, or 0.50 or larger, respectively, according to the interpretation of effect sizes proposed by Cohen (2013). Since effects of the driving course should rather appear on relative (than on absolute) power values (see above), the correlation analysis was exclusively performed for alpha and theta fractions.

The track-frequency analysis was completed by a simple classifying algorithm intended to estimate the track-specific task load based on the EEG data. Here, it was assumed that high theta activity is associated with increased mental effort and high alpha activity with reduced attentional allocation. Therefore, as in our previous analysis (Wascher et al., 2019), the algorithm tested theta and alpha fraction against each other using a paired-sample *t*-test for each data point (i.e., for each 10-meter track segment), and assigned *low task load* to track segments with significantly higher alpha than theta fraction, and *high task load* to segments with significantly higher theta than alpha fraction. If theta and alpha fraction did not differ significantly, a *median task load* was assigned to this track segment. This classification procedure was performed equally for MP1 and MP2, and it was determined how many road sections were rated as equally difficult at both measurement times (suggesting a reliable estimation of task load from the EEG) or were rated as easier or more difficult in MP1 and MP2. Finally, for each participant the correlation of EEG-based task load estimates at MP1 and MP2 was computed across the entire 30-km driving course (i.e., across the each 10-meter track segment), using Pearson's r correlations.

RESULTS

Behavior

The track-based analyses of the driving parameters showed that both driving speed and steering wheel angular velocity profoundly varied along the driving course (**Figure 2**). In particular, while the freeway section (FW) was characterized by high driving speed and low steering angular velocity (apart from passing through a freeway parking area at kilometer 15), the second state road section (SR2) and especially the city traffic drive (CT) were characterized by lower and highly varying driving speed as well as increased and higher steering angular velocities. More importantly, however, it is to notice that the driving speed increased overall, while the steering wheel angular velocity decreased at MP2 relative to MP1. Also, the mean drive time for the entire course went from 51.6 min (SD 9.0) to 46.7 min (SD 8.3), $t_{(126)} = 5.65$, $p < 0.001$.

These differences were even more evident in the task-load related analysis, analyzing the driving parameters separately for passages with simple, complex, and interactive driving profiles (**Figure 3**). The mean driving speed significantly increased from MP1 to MP2, $F_{(1, 126)} = 25.60$, $p < 0.001$, adj $\eta^2_p = 0.162$, while the mean steering wheel angular velocity decreased, $F_{(1, 126)} = 20.15$, $p < 0.001$, adj $\eta^2_p = 0.131$. There were no interactions of measurement time and driving profile, neither for driving speed, $F_{(2, 252)} = 0.35$, $p = 0.71$, adj $\eta^2_p = 0.005$, nor for steering wheel angular velocity, $F_{(2, 252)} = 1.36$, $p = 0.26$, adj $\eta^2_p = 0.003$, indicating that the effects of driving profile on driving speed, $F_{(2, 252)} = 4{,}402.95$, $p < 0.001$, adj $\eta^2_p = 0.972$, and steering angular velocity, $F_{(2, 252)} = 2{,}118.03$, $p < 0.001$, adj $\eta^2_p = 0.944$, did not depend on measurement time. Thus, the participants drove at highest speed in simple passages and significantly reduced the speed in complex passages, $F_{(1, 126)} = 5{,}488.40$, $p < 0.001$, adj $\eta^2_p = 0.977$. Relative to complex passages, they also drove faster when there were interactions with other road users, $F_{(1, 126)} = 1{,}369.34$, $p < 0.001$, adj $\eta^2_p = 0.915$. The steering angular velocity increased from simple to complex passages, $F_{(1, 126)} = 5{,}919.92$, $p < 0.001$, adj $\eta^2_p = 0.979$, and further from complex to interactive passages, $F_{(1, 126)} = 169.92$, $p < 0.001$, adj $\eta^2_p = 0.571$.

Alpha and Theta Power Analysis

The track-based analysis of brain oscillatory power demonstrated that both alpha and theta power fractions varied substantially over the driving route (**Figure 4**): Phases of high alpha fraction alternated with short sections in which alpha fraction was strongly reduced. For example, the freeway passage was characterized by high alpha fraction values, while these were reduced at the beginning of the fog passage at kilometer 3, when driving through the freeway parking area at kilometer 15, and during city driving after kilometer 27. The theta values, on the other hand, showed a rather inverse pattern.

The task-load related analysis indicated that raw alpha power significantly increased from MP1 to MP2, $F_{(1, 126)} = 10.02$, $p < 0.005$, adj $\eta^2_p = 0.066$, while differences for raw theta power and alpha and theta fraction power were not significant, all *F*s < 2.13, all *p*s > 0.14 (**Figure 5**). There were effects of driving profile on raw theta power, $F_{(2, 252)} = 24.75$, $p < 0.001$, adj $\eta^2_p = 0.158$, as well as alpha fraction power, $F_{(2, 252)} = 11.89$, $p < 0.001$, adj $\eta^2_p = 0.079$, and theta fraction power, $F_{(2, 252)} = 66.38$, $p < 0.001$, adj $\eta^2_p = 0.340$, but not raw alpha power, $F_{(2, 252)} = 0.15$, $p = 0.86$, adj $\eta^2_p = 0.007$. Also, there were no interactions of measurement time and driving profile, all *F*s < 2.76, all *p*s > 0.06. Further comparisons of the different driving profiles indicated that alpha fraction decreased from simple to complex passages,

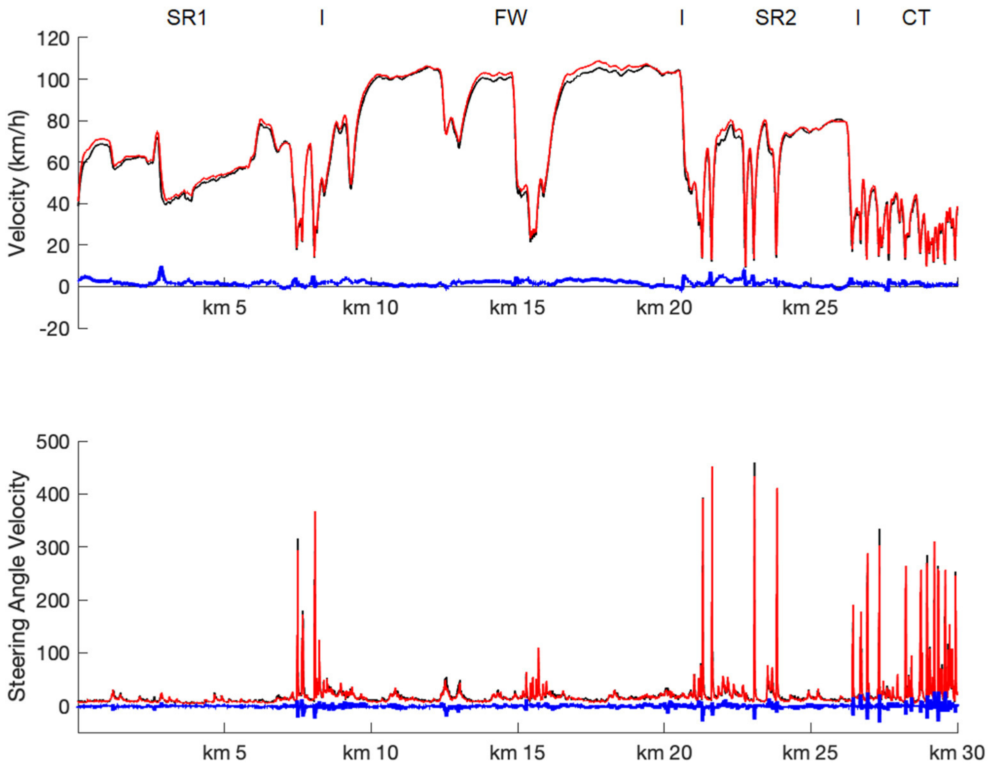


FIGURE 2 | Track-based analyses of driving parameter. Mean driving speed (upper row) and steering wheel angular velocity (lower row) as function of driving route, shown separately for MP1 (black), MP2 (red), and MP1 – MP2 differences (blue). Note that for each time point individual values of each participant were assigned to fix waypoints and then averaged 10-meter wise. SR1, first state road; FW, freeway; SR2, second state road; CT, city traffic.

$F_{(1, 126)} = 14.36$, $p < 0.001$, adj $\eta_p^2 = 0.095$, but did not differ in complex and interactive passages, $F_{(1, 126)} = 0.06$, $p = 0.82$, adj $\eta_p^2 = 0.007$. Raw theta power increased from simple to complex passages, $F_{(1, 126)} = 44.27$, $p < 0.001$, adj $\eta_p^2 = 0.254$, but did not differ in complex and interactive passages, $F_{(1, 126)} = 0.04$; $p = 0.85$, adj $\eta_p^2 = 0.008$. Theta fraction power also increased from simple to complex passages, $F_{(1, 126)} = 52.53$, $p < 0.001$, adj $\eta_p^2 = 0.289$, and was stronger in interactive than in complex passages, $F_{(1, 126)} = 10.95$, $p < 0.005$, adj $\eta_p^2 = 0.073$.

Correlational Alpha and Theta Power Analysis

In order to estimate the degree to which alpha and theta fraction power remained stable between the two measurement points at an intraindividual level, correlations have been computed across the entire 30-km driving course (i.e., across the 3,000 10-meter track segments) for each participant. Individual analyses revealed that the correlation coefficients were quite evenly distributed and ranged from low to medium (**Figure 6**). There were significant positive correlations ($p < 0.05$) in 73.2% of the participants for alpha fraction, $r = 0.57$–0.04, and in 80.3% for theta fraction, $r = 0.59$–0.04. Of these significant positive correlations, 53.8% (alpha fraction) and 68.6% (theta fraction) were in a low range, $r > 0.1$, and 7.5% (alpha fraction) and 13.7% (theta fraction) were in a medium range, $r > 0.3$. Also, significant negative correlations were found in 6.3% of the participants for alpha fraction, $r = -0.10$ to -0.04, and in 6.3% for theta fraction, $r = -0.13$ to -0.04.

EEG-Based Estimation of Task Load

The EEG-based estimation of the track-specific task load revealed a pronounced variance of load ratings along the driving course (**Figure 7**). High load ratings were mainly found at the beginning of the drive and of the fog passage (at kilometer 3), during the state road sections (SR1 and SR2) as well as during the city traffic drive (CT). Low load ratings were found during the freeway section (FW), apart from passing through a freeway parking area (at kilometer 15). This pattern was overall quite similar at MP1 and MP2, $r = 0.729$. There were, however, some differences in task load ratings: Higher ratings were found at the beginning and the end of the fog passage (at kilometer 3 and 5), while passing through the freeway parking area (at kilometer 15), and at the end of the freeway section. In contrast, lower ratings were found during the fog passage, during the freeway section (FW), and the second state road section (SR2). Overall, of the 3,000 (10-meter) track segments assessed, 72.23% were rated the same in terms of task load, 14.66 % were rated as easier and 13.11% were rated as more difficult. Not a single road section that was rated as easy (difficult) in one of the two measurements was rated as difficult (easy) in the other measurement (**Figure 8**).

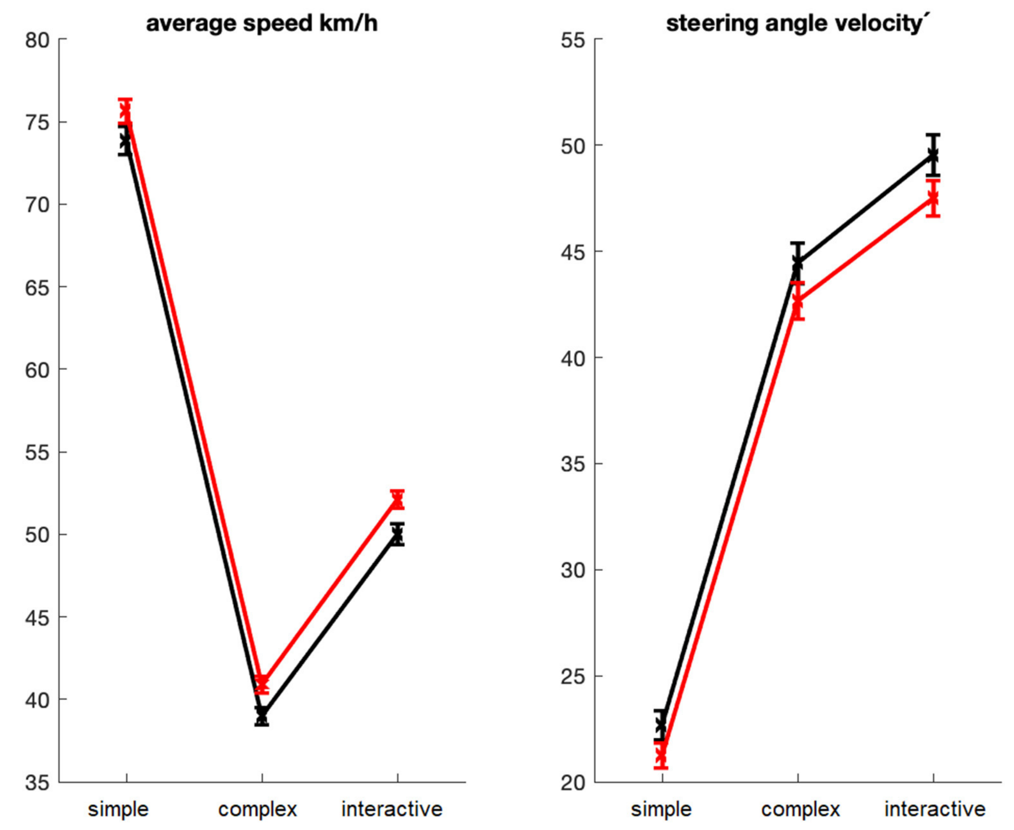


FIGURE 3 | Task-load related analysis of driving parameters. Average driving speed (left) and steering angle velocity (right) as function of driving profile (simple, complex, interactive), shown separately for MP1 (black) and MP2 (red). Error bars indicate standard errors.

Finally, in order to test to what extent the EEG-based estimates of task load are consistent at the first and second measurement time points at an intraindividual level, correlations were computed across the 3,000 (10-meter) track segments for each participant. There were significant positive correlations in 81.9% of the participants, $r = 0.44–0.04$, of which 61.5% were in a low range, $r > 0.1$, and 4.8% in a medium range, $r > 0.3$. Significant negative correlations occurred in 7.1% of the participants, $r = -0.11$ to -0.04 (all $p < 0.05$).

DISCUSSION

The aim of the present study was to evaluate the reliability of the cEEGrid technology in a longitudinal investigation of the driving abilities of older adults. Behavioral and electrophysiological parameters of mental load measured while driving in a driving simulator at two time points more than 1 year apart were compared and related to characteristics of the driving course. With a high reliability of the measurement, comparable effects of task difficulty on the EEG parameters should appear (independent of the time of measurement), which should also be related to the behavioral measures. In addition, a high correspondence of the oscillatory measures between the first and the second measurement time should occur on an intra-individual level. The analyses indicated a number of specific effects of measurement time point and driving profile on behavioral driving parameters and brain oscillatory activity that are discussed in detail in the following.

Driving Parameters: Speed and Steering Wheel Angular Velocity

Overall, the average speed increased while the steering wheel angular velocity decreased from MP1 to MP2. Given that the driving speed in our scenario could be freely chosen by the driver within the maximum speed limits, the increase in driving speed at the second measurement time point could indicate an increase in perceived safety when managing the driving task at a second time. On the one hand, this could result from a higher familiarity with the route. Especially in elderly drivers, a reduction of speed is a frequently observed strategy when driving an unknown route or when the driving situation becomes more complex so that drivers feel unsafe (Trick et al., 2010). In extreme cases, this can lead to dangerous driving situations, for example, if other road users are hindered and forced to make unnecessary and risky overtaking maneuvers. On the other hand, driving speed is usually increased with decreasing workload (Harms, 1986; Verwey and Veltman, 1996), which would also suggest that the second drive was less challenging to the participants than the first one.

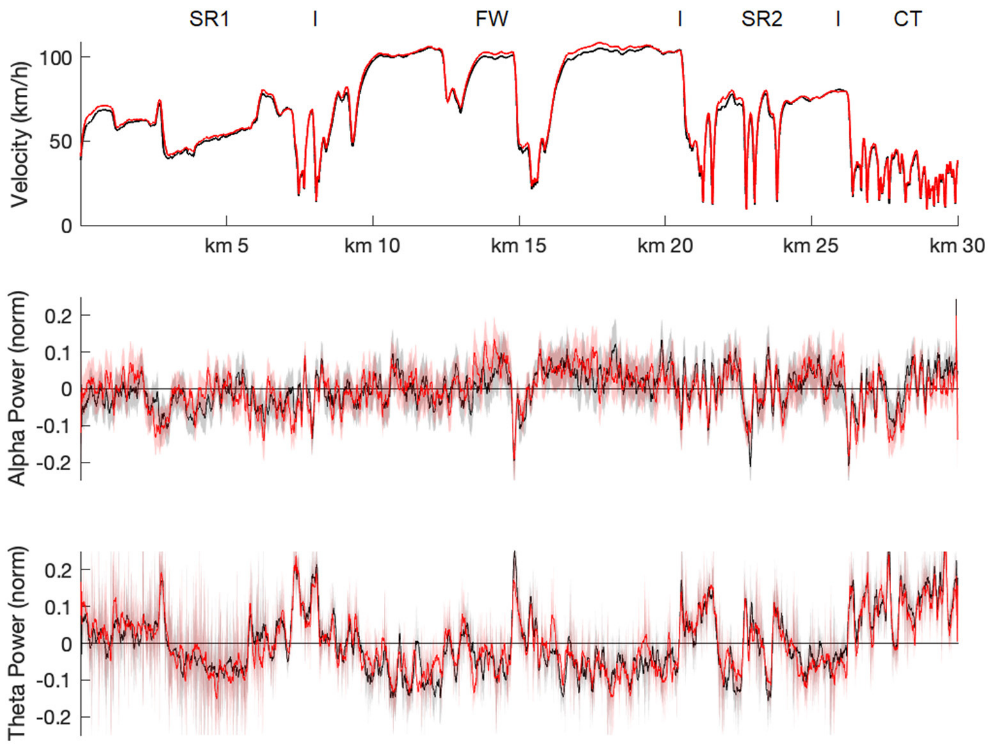


FIGURE 4 | Track-based analyses of EEG parameter. Mean z-transformed alpha (middle row) and theta (lower row) power fractions as function of driving route for MP1 (black) and MP2 (red), shaded by their 95% confidence intervals. For comparison, mean driving speed (upper row) is also shown. Note that for each time point individual alpha and theta values of each participant were assigned to fix waypoints and then averaged separately for 10-meter track segments. SR1, first state road; FW, freeway; SR2, second state road; CT, city traffic.

This is also supported by the decrease in steering wheel angular velocity as this measure is also considered to be an indicator of task load while driving (e.g., Antin et al., 1990; Verwey and Veltman, 1996). Here, high angular velocities are associated with high load whereas low angular velocities are associated with low load. Accordingly, repeatedly driving the same route would be less stressful than maneuvering on a completely unfamiliar route. It is remarkable that there was more than 1 year between the measurements, which means that the participants seem to have memorized the requirements of the route very well. In addition, long-term learning effects could play a role by which the participants benefit from a more and more experienced anticipation of steering behavior of the car. In a previous driving simulator study, in which younger and older participants had to keep a virtual car on track on a curvy road, we also observed learning effects in form of a decrease in steering variability during the ∼1-h drive (Getzmann et al., 2018).

With regard to the reliability of the measurements, it is also remarkable that the influences of the driving profile on driving speed and steering wheel angular velocity did not differ at MP1 and MP2. The driving course was subdivided into simple, complex, and interactive driving profiles, which were related to different levels of task load, based on known factors of mental load in driving situations (Pauzié, 2008; Engström et al., 2017; Rahman et al., 2017). Thus, passages with an undisturbed ride on a free route were rated as of low task load, passages with junctions with turning, roundabouts, and left turns as of medium task load, and interactions with other traffic participants as of high task load (for a critical discussion, see Wascher et al., 2019). The increase in steering angular velocity with increasing task load corresponds well with the assumption that this measure is associated with the demands of driving, which, as expected, is lower for a simple driving profile than for a complex one (involving intersections and traffic lights) as well as interactions with other road users. A limiting factor here could be that the driving profiles were not evenly distributed over the route sections. For example, complex driving profiles (with intersections and give way signs) are more common in the city, while freeway sections are more characterized by simple driving profiles. This might also explain the (unexpectedly) higher driving speed with interactive than complex driving profiles: Interactions with other road users are also common on state roads and freeways (where driving speed is on average higher than in the city), whereas complex driving profiles (and lower driving speed) are more common in the city. An interaction of driving profile and route section can therefore not be completely ruled out. Nevertheless, the replication of this general pattern suggests that overall demands decreased relative to the first test drive, but did not depend on the route profiles passed through.

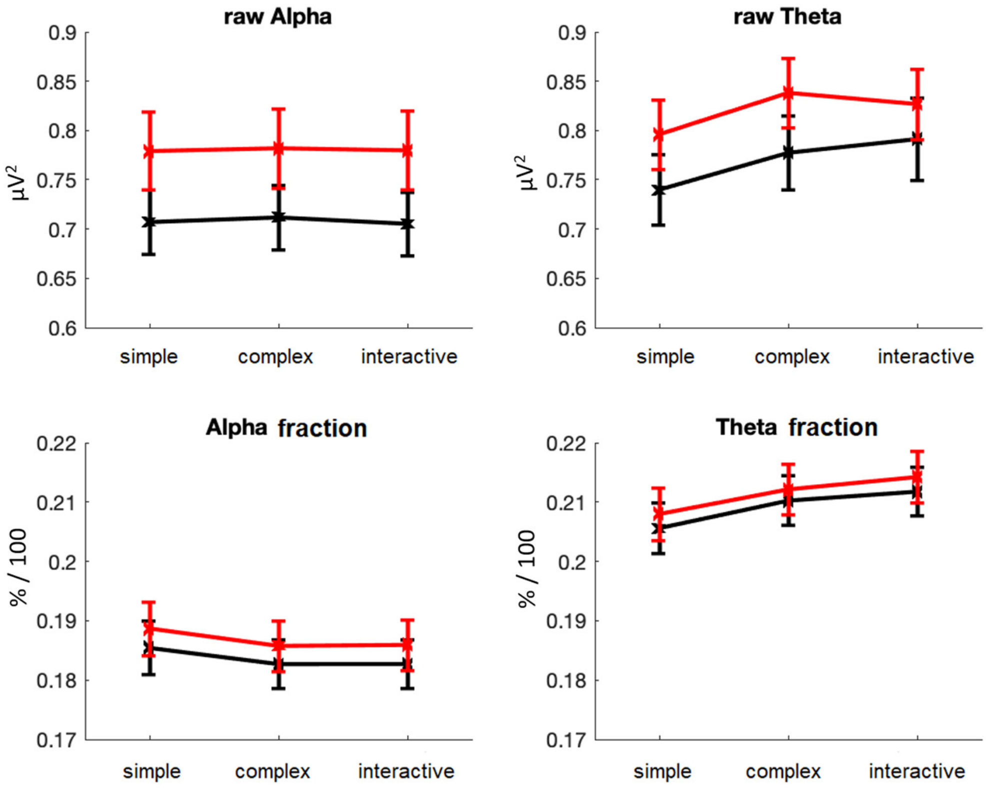


FIGURE 5 | Task-load related analysis of EEG parameters. Raw (upper row) and fractional (lower row) average power in the alpha (left) and theta (right) frequency bands as function of driving profile (simple, complex, interactive), shown separately for MP1 (black) and MP2 (red). Error bars indicate standard errors.

EEG Parameters: Alpha and Theta Power

Comparable patterns to the behavioral data were also found in the derived EEG measures: Overall, raw alpha activity increases from the first to the second measurement time point. In addition, and independently of the time of measurement, relative alpha power (power fraction) varied with the driving profiles and was higher at simple compared to complex and interactive passages. In general, decreases in alpha power are usually associated with the allocation of attention (Herrmann and Knight, 2001), while increases in alpha power is assumed to reflect mental fatigue, but also attentional withdrawal and disengagement (Hanslmayr et al., 2012; Wascher et al., 2014, 2016) as typically observed when tasks are perceived as monotonous and boring (Borghini et al., 2014). In the driving context, increased alpha power has thus been observed during monotonous driving situations, probably reflecting periods of inattention and mind-wandering (Lin et al., 2016). Assuming increases in alpha activity to be associated with reduced attentional engagement, the present findings would argue for a withdrawal of attentional resources, both in longitudinal and route-related terms: The participants seemed to pay less attention to the driving task when they drove the same route for a second time. However, they continued to flexibly adapt their mental resources to the task demands and increased their attention when the traffic situation became more complex. Interestingly, effects of driving profile were only found on alpha fraction power, but not raw alpha power. This discrepancy could be due to a relatively high power in oscillatory activity in low frequency bands (as has been observed in Wascher et al., 2019), which could have masked experimental effects in the higher frequency alpha band. In line with this assumption, it has been shown that weak effects in higher frequency bands tend to become evident in relative power measures rather than in absolute (raw) measures, where low-frequency power is dominant (Labounek et al., 2019).

Activity in theta power is generally associated with cognitive control (Cavanagh and Frank, 2014; Cavanagh and Shackman, 2015) and typically increased with higher workload (Wilson and Hankins, 1994; Gevins et al., 1997) and task demands (Lal and Craig, 2001; Jensen and Tesche, 2002; Onton et al., 2005; Borghini et al., 2014). Theta power also increases with higher task engagement (Yamada, 1998; Onton et al., 2005) and with the effort to keep task performance high (Wascher et al., 2014; Arnau et al., 2017). In line with this assumption, both raw and relative (fractional) theta power values were increased in the present driving task at more complex route sections, such as at the beginning of the fog passage and during city driving. However, independently of these demand-related modulations, there was rather an (albeit not significant) increase in raw theta power at the second measurement time point (cf. **Figure 4**), suggesting that

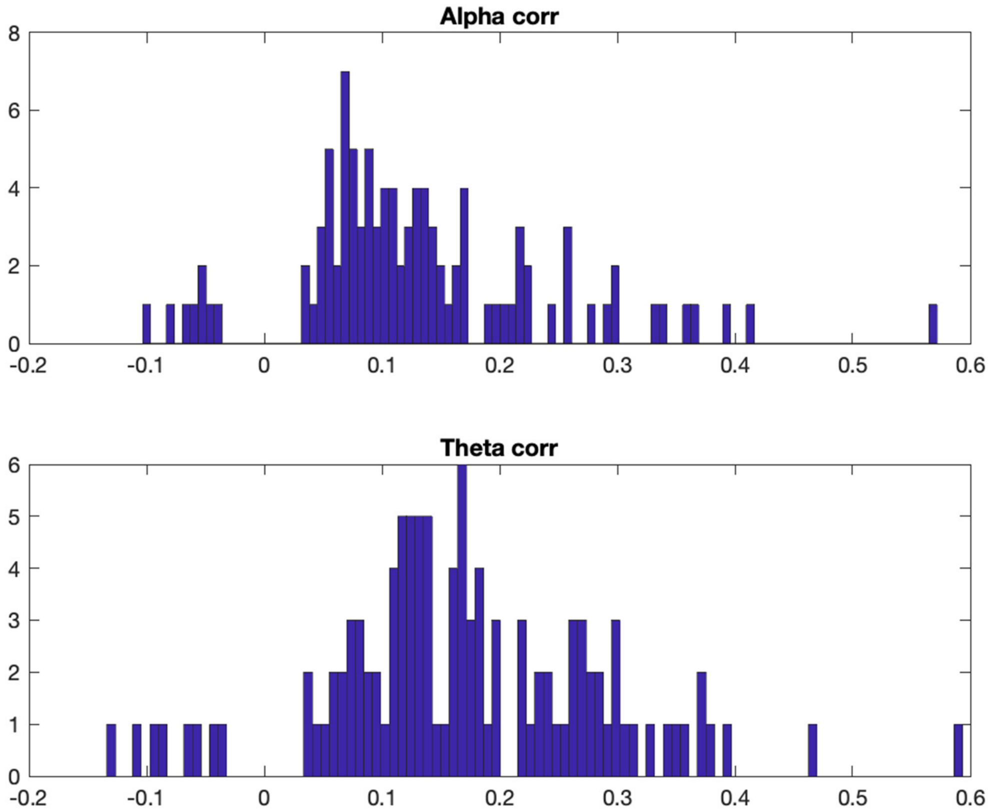


FIGURE 6 | Frequency distribution of individual (significant) Pearson's correlation coefficients for alpha (upper row) and theta (lower row) fraction power measured at MP1 and MP2.

the task demands and/or task engagement increased (or at least did not differ) at MP2 compared to MP1.

One could argue that the theta power findings partly contradict the interpretation of the alpha power, suggesting a decrease in task engagement. Selection effects in the way that drivers with low task engagement left the study after MP1 and did not participate in MP2 can be excluded, as the same drivers were examined at both measurement time points. A more plausible interpretation could be that the participants had a higher motivation to perform well in the driving task, perhaps even better than at the first time. Given that the current study is designed to detect age-related deteriorations in driving ability, the participants' motivation to counteract these by increasing effort may be particularly pronounced, as reflected by undiminished theta activity. This interpretation is supported by findings of a previous study on age-related differences in pro-active driving behavior (Getzmann et al., 2018): Better performance in proactive driving (i.e., more alert steering behavior, better anticipation and active use of driving-relevant information and more proactive planning of driving manoeuvers) was associated with increased mental effort in the older group, as reflected by higher theta power. Moreover, only in the older group a relationship between steering variability and theta power was found, with better steering performance being associated with higher theta power. Taken together, the EEG findings suggest that the drivers were more relaxed, but remained motivated to perform the driving task well at the second time.

Another relevant aspect to be discussed here are task-specific differences between alpha and theta activity, which are also reflected in differences in the brain areas over which they are usually derived. While alpha power is most prominent over occipital-parietal areas of visual cortex, theta power is measured over frontal areas associated with higher cognitive executive functions (for review, Klimesch, 1999). In a realistic driving task, in which complex and monotonous driving passages alternate, and in which multiple subtasks such as visual information uptaking and processing, attention allocation, spatial navigation have to be performed, alpha and theta activity should therefore be differently involved (Di Flumeri et al., 2018; Puma et al., 2018; Wang et al., 2018). In particular, alpha activity (i.e., its suppression) seems to be rather associated with task engagement, while theta activity seems to be associated with task workload (Berka et al., 2007; Wang et al., 2018). This could explain, for example, differences in the dependence of alpha and theta power on the driving profile. For example, the track-based analysis indicated an increase in theta activity at the beginning of the second state road section at kilometer 21 (which was characterized by demanding passages), which was not accompanied by a suppression of alpha activity (cf. **Figure 4**). Also, theta fraction power was higher in interactive than in

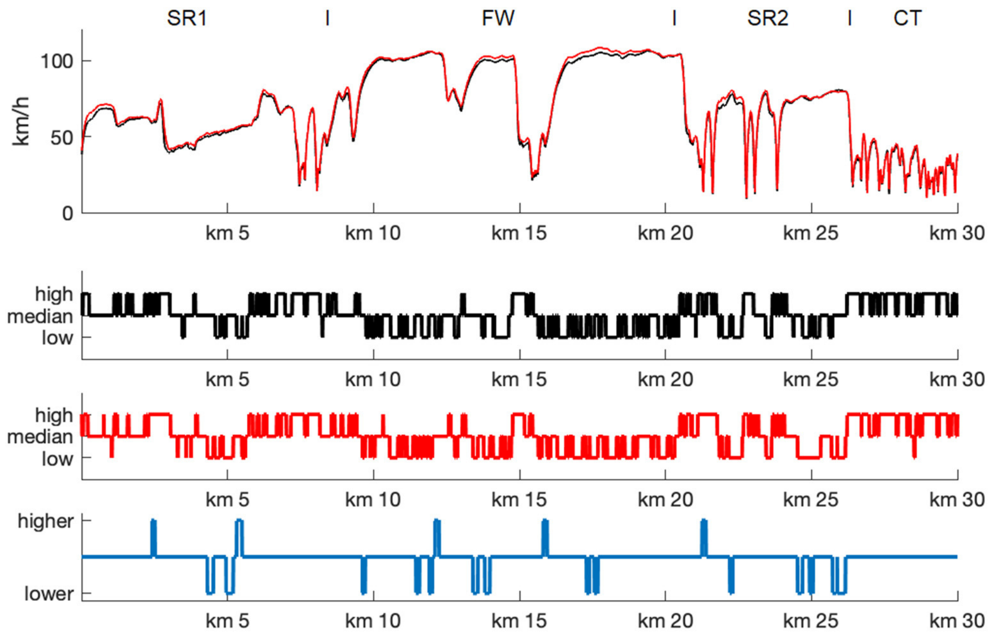


FIGURE 7 | EEG-based estimation of task load. Task load (low, medium, high) as function of driving route for MP1 (black) and MP2 (red), and MP1 – MP2 differences (blue). For each time point individual task load estimations of each participant were assigned to fix waypoints and then averaged separately for 10-meter track segments. For comparison, mean driving speed (upper row) is also shown. SR1, first state road; FW, freeway; SR2, second state road; CT, city traffic.

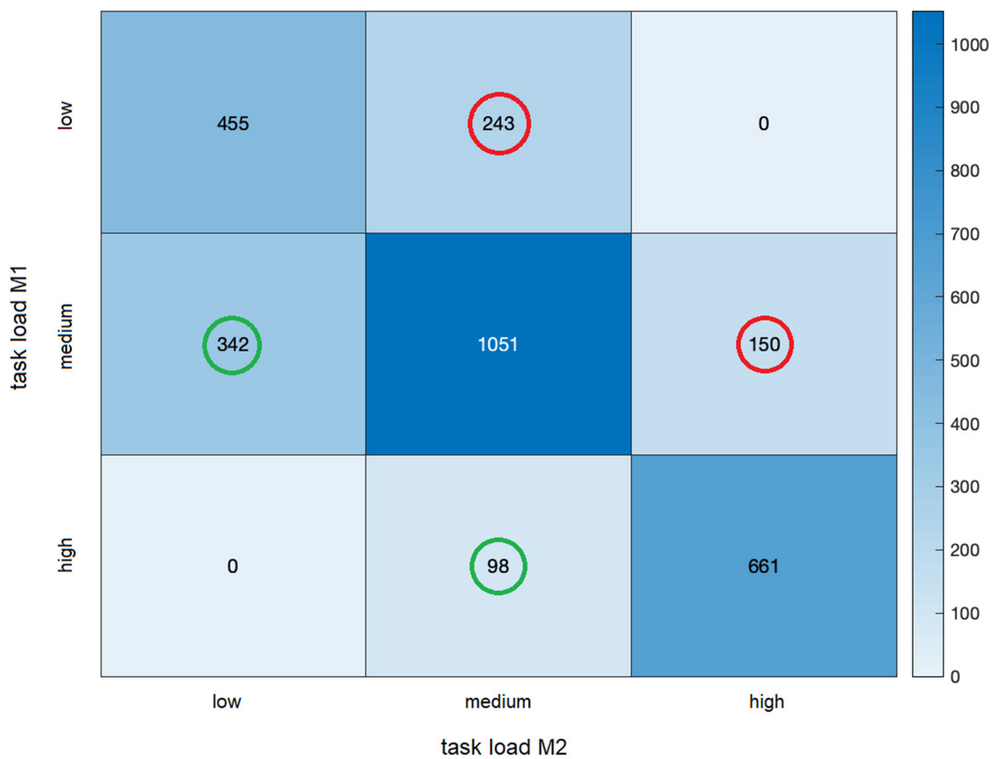


FIGURE 8 | Heatmap of EEG-based estimation of task load at MP1 and MP2. The figure shows the number of road sections that were rated as equally difficult at both measurement times, as well as the number of sections that were rated as easier (outlined in green) or more difficult (outlined in red) in MP2 than in MP1, averaged across all participants. A total of 3,000 (10-meter) track segments were classified as of low, medium, or high task load.

complex driving profiles, whereas alpha fraction power did not differ in either condition. This could suggest that interaction with other road users and driving on a demanding but empty route account for differences in task load, while task engagement is hardly affected.

A limiting factor which complicates the interpretation of the results, and which has to be noted here is the reduced spatial resolution of the cEEGrid technology. The electrodes are located largely over temporal areas, making a clear differentiation of oscillatory activity into frontal and parieto-occipital parts difficult. Traditionally, theta activity is derived over frontal areas and alpha activity over posterior areas, where the power in these frequency bands is usually most prominent (for review, Klimesch, 1999). Thus, only very few studies have considered theta and alpha activity measured over temporal areas as potential indices of mental workload and task engagement. In studies on simulated driving (Diaz-Piedra et al., 2020) and multi-tasking (Puma et al., 2018), workload-induced modulations of theta activity (with higher workload being associated with higher theta activity) were not only observed over frontal and occipital regions, but also over temporal regions. A combined EEG-fMRI study showed that workload-induced modulations of theta activity were most pronounced over frontal and posterior areas (Sammer et al., 2007). However, an additional EEG-constrained fMRI analysis revealed that the generators of these effects were not primarily localized frontally, but form a network including temporal and hippocampal hemodynamic activation, cingulate activation, frontal superior, and cerebellar activation. The authors thus concluded that theta band activity reflects a binding process of widely distributed cortical areas, which all contribute to the EEG activity derived at the scalp. The same could be true for alpha band activity, which appears to reflect a network-binding mechanism, supporting the interplay within thalamo-cortical networks relevant for sensory gating and the control of vigilance and attention (Lopes da Silva, 2013; for review, Nishida et al., 2015). Significant effects of task difficulty on alpha power (with easier task conditions being associated with larger power) have been observed over temporal areas (Brookings et al., 1996), while other studies failed to find effects of performance (Çiçek and Nalçaci, 2001) and relaxation (Scholz et al., 2018) on alpha activity over temporal areas, which were observed over parietal areas. Thus, it appears that theta effects could be more reliably derived over temporal areas than alpha effects. This could also explain why in the present study (as well as in our previous study, Wascher et al., 2019) discrepancies between raw and fractional power occurred in the alpha band, but not in the theta band: Given that fractional power was corrected for total power in the signal, alpha effects could be more pronounced (independent of their topography) in fractional power. However, since no conventional multi-channel EEG cap has been employed here for a direct comparison of the signals measured with cEEGrids, especially the interpretation of the alpha activity should be treated with caution.

Test-Retest Reliability Considerations

Two different approaches have been chosen to determine the retest reliability of the EEG results, first a task-load related analysis, investigating whether the EEG measures at both measurement points depended on the driving profile in the same way, and second an (intra-individual) correlational analysis, comparing the EEG measures along the route (subdivided into 3,000 10-m track segments) at MP1 and MP2 separately for each subject. The task-load related analysis showed a high correspondence of the EEG patterns between the two measurement times across all participants: That is, independent of the measurement time, challenging traffic situations are accompanied by a reduction of alpha and an increase of theta (e.g., as can be seen at the beginning of the city drive), whereas monotonous traffic situations (e.g., the foggy passage or the undisturbed highway drive) showed the opposite pattern. Thus, the effect of driving profile on alpha and theta activity was reliably found at both measurement times, indicating a high reliability of the measurement, especially for fraction values.

The same is true for the high correspondence in the estimation of task load from the alpha and theta values between the two measurement time points. The track-based analysis indicated that passages that were estimated to be easy (or hard) at MP1 were also easy (or hard) at MP2. In particular, averaged across all participants, the analysis showed that not a single road section that was rated as easy (difficult) in MP1 or MP2 was rated as difficult (easy) in the other measurement (cf. heatmap in **Figure 8**), which indicates a reliable estimation of task load. On the other hand, this also means that road sections overall were not estimated to be easier at MP2. Thus, a higher familiarity with the route (suggested by a higher average speed and lower steering wheel angular velocity) was not associated with a reduced difficulty (estimated from the alpha and theta ratio). In other words: a difficult passage (associated with high theta and low alpha activity) may well be passed more quickly due to familiarity with the route, without it becoming less challenging. Still, a few changes emerged that can be plausibly explained (as can the driving parameters). For example, the patterns of alpha/theta values at the second measurement time point indicate an increased task load at the beginning and end of the fog passage, whereas during the fog passage the task load was estimated to be lower. Both effects can be explained by an increasing familiarity of the participants with the route. This interpretation is in line with the so-called "route-familiarity effect," in which greater route familiarity can lead to increased inattention and mind-wandering and, as a consequence, to driving impairments (e.g., Martens and Fox, 2007; Yanko and Spalek, 2013). The same was true for undisturbed country road passages, which appeared to be driven with a higher routine and lower task load. This and the overall high correspondence of the EEG patterns with the behavioral data suggest high content test-retest reliability of the cEEGrids technology used for the sample of participants as a whole.

On an intraindividual level, significant positive correlations were found for most of the participants, both for alpha and theta activity as well as for the derived EEG-based estimation of task load. Participants who showed a high alpha or theta activity at the first measurement time and a high mental workload did so again at the second measurement time, which indicates a certain degree of temporal stability of the measurements. However, it has to be noted that the vast majority of the individual correlations were (although mostly highly significant) in a low range. Usually, higher reliabilities are found in more structured EEG conditions (McEvoy et al., 2000), i.e., in demanding cognitive tasks, since fluctuations in cognitively engaging tasks are generally lower. At least in complex road sections and in interaction with other road users, an increased cognitive load can also be assumed.

Regarding an individual (possibly diagnostic) evaluation, the small correlations make the interpretation in terms of a change from MP 1 to MP 2 difficult. The same applies to the prediction of future values based on previous values. In addition, it is remarkable that some (few) participants also showed negative correlations, suggesting that the pattern of oscillatory power over the driving distance has (at least partly) reversed. This could indicate a change in the mental resources that some subjects invested in the driving task, with a high task engagement at MP1 changing to an attentional disengagement at MP2 (or vice versa). In this context, it should be pointed out that the data come from an ongoing study on the development of traffic safety parameters in older drivers, and that changes in mental abilities are to be expected in the age range considered.

In summary, however, it must be stated that the correlations within the participants are rather low, i.e., that the alpha/theta activity in track segments at MP1 is poorly associated with the alpha/theta activity in the same segment at MP2. This suggests high fluctuations in oscillatory activity between measurement time points that are not related to the task load of the track segments themselves. It is difficult to assess whether this is due, for example, to transient fluctuations in mental states like alertness and vigilance during the drive, or changes within participants over the relatively long time period between MP1 and MP2, or demonstrates limitations of the EEG methodology used. Further insights may be provided by the investigation of possible correlations between changes in individual driving performance (and its changes over time) and EEG parameters, which are planned at a later stage of the still ongoing project. The age range of the test group, which is clearly not representative for the entire population of drivers, may also be a potentially limiting factor with regard to the generalizability of the results. Age-related decreases in cognitive performance as well as increases in interindividual variation both can lead to a conflict with the determination of the reliability of the EEG method. Another problem specific to the cEEGrid technology is that the electrodes on older skin, which is often drier and more wrinkled, may have increased resistances, resulting in poorer and fluctuating conduction of the EEG. Future comparative studies with younger subjects therefore seem appropriate.

CONCLUSIONS

Taken together, the present test-retest analysis demonstrated changes in behavioral and brain oscillatory parameters between the first and second measurement time point across all participants, which can be characterized by an increase in driving speed and decrease in steering angular velocity as well as an increase in alpha power, while theta power remained rather stable. These changes suggest a reduced overall task load which appears plausible with regards to learning and memory effects. At both measurement points, the EEG parameters (like the behavioral parameters) were similarly modulated by track difficulty and—as a consequence—task demands, indicating a high reliability and ecological validity of the EEG application via cEEGrid technology. At the intra-individual level, positive correlations of the oscillatory measures and its dependence on track difficulty were found in the majority of the participants. On the other hand, intra-individual correlations were (although significant) rather low, raising the question of the individual-diagnostic value of the chosen method. Further analysis of the reasons why some participants showed significant differences compared to the first measurement will be necessary to determine if this was due to the EEG recording or if the causes may be found in the participants themselves (e.g., cognitive decline). However, in the context of task-related EEG parameters which maps changes in cognitive states related to, for example, task difficulty, the reliability and ecological validity of cEEGrid electrodes appear satisfactory. Overall and in combination with the findings of our previous study (Wascher et al., 2019), the results provide further evidence for the usability of portable low-density EEG methods like cEEGrids as an alternative to conventional lab-based recording systems for mapping mental processes in natural environments.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the funder of this study does not allow sharing of data. Requests to access the datasets should be directed to Stephan Getzmann, getzmann@ifado.de.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Local Ethics Committee of the Leibniz Research Centre for Working Environment and Human Factors. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SG, GR, MK, and EW were responsible for the design of the study. EW analyzed the data. JR was involved in data acquisition and provided data and analyses of a pilot-study. SG wrote a first version of the manuscript. All authors contributed to the completion of the final version of the manuscript.

FUNDING

This report is based on parts of the research project carried out at the request of the Federal Ministry of Transport and Digital Infrastructure, represented by the Federal Highway Research Institute, under research project No. 82.0649/2016. The authors are solely responsible for the content. The publication of this article was funded by the Open Access Fund of the Leibniz Association.

ACKNOWLEDGMENTS

The authors are grateful to Tobias Blanke and Trung Tin Pham for preparing the software and parts of the electronic equipment, to Virginia Herman and her team of student assistants for their help in running the experiment, and to three reviewers for very constructive comments on a previous version of the manuscript.

REFERENCES

Angelidis, A., van der Does, W., Schakel, L., and Putman, P. (2016). Frontal EEG theta/beta ratio as an electrophysiological marker for attentional control and its test-retest reliability. *Biol. Psychol.* 121, 49–52. doi: 10.1016/j.biopsycho.2016.09.008

Antin, J. F., Dingus, T. A., Hulse, M. C., and Wierwille, W. W. (1990). An evaluation of the effectiveness and efficiency of an automobile moving-map navigational display. *Int. J. Man Mach. Stud.* 33, 581–594. doi: 10.1016/S0020-7373(05)80054-9

Arnau, S., Möckel, T., Rinkenauer, G., and Wascher, E. (2017). The interconnection of mental fatigue and aging. An EEG study. *Int. J. Psychophysiol.* 117, 17–25. doi: 10.1016/j.ijpsycho.2017.04.003

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., et al. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviat. Space Environ. Med.* 78, B231–244.

Bleichner, M. G., and Debener, S. (2017). Concealed, unobtrusive ear-centered EEG acquisition: cEEGrids for transparnt EEG. *Front. Hum. Neurosci.* 11, 1–14. doi: 10.3389/fnhum.2017.00163

Bleichner, M. G., Mirkovic, B., and Debener, S. (2016). Identifying auditory attention with ear-EEG: cEEGrid versus high-density cap-EEG comparison. *J. Neural Eng.* 13:066004. doi: 10.1088/1741-2560/13/6/066004

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., and Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neurosci. Biobehav. Rev.* 44, 58–75. doi: 10.1016/j.neubiorev.2012.10.003

Brookings, J. B., Wilson, G. F., and Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biol. Psychol.* 42, 361–377. doi: 10.1016/0301-0511(95)05167-8

Cavanagh, J. F., and Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* 18, 414–421. doi: 10.1016/j.tics.2014.04.012

Cavanagh, J. F., and Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. *J. Physiol.* 109, 3–15. doi: 10.1016/j.jphysparis.2014.04.003

Chang, C. Y., Hsu, S. H., Pion-Tonachini, L., and Jung, T. P. (2019). Evaluation of artifact subspace reconstruction for automatic artifact components removal in multi-channel EEG recordings. *IEEE Trans. Biomed. Eng.* 67, 1114–1121. doi: 10.1109/TBME.2019.2930186

Chiang, A. K. I., Rennie, C. J., Robinson, P. A., Van Albada, S. J., and Kerr, C. C. (2011). Age trends and sex differences of alpha rhythms including split alpha peaks. *Clin. Neurophysiol.* 122, 1505–1517. doi: 10.1016/j.clinph.2011.01.040

Çiçek, M., and Nalçaci, E. (2001). Interhemispheric asymmetry of EEG alpha activity at rest and during the Wisconsin Card Sorting Test: relations with performance. *Biol. Psychol.* 58, 75–88. doi: 10.1016/S0301-0511(01)00103-X

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Academic Press.

Debener, S., Emkes, R., De Vos, M., and Bleichner, M. (2015). Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear. *Sci. Rep.* 5:16743. doi: 10.1038/srep16743

Di Flumeri, G., Aric,ò, P., Borghini, G., Sciaraffa, N., Di Florio, A., and Babiloni, F. (2019). The dry revolution: Evaluation of three different EEG dry electrode types in terms of signal spectral features, mental states classification and usability. *Sensors* 19:1365. doi: 10.3390/s19061365

Di Flumeri, G., Borghini, G., Aric,ò, P., Sciaraffa, N., Lanzi, P., Pozzi, S., et al. (2018). EEG-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings. *Front. Hum. Neurosci.* 12:509. doi: 10.3389/fnhum.2018.00509

Diaz-Piedra, C., Sebastián, M. V., and Di Stasi, L. L. (2020). EEG theta power activity reflects workload among army combat drivers: an experimental study. *Brain Sci.* 10:199. doi: 10.3390/brainsci10040199

Engel, A. K., Maye, A., Kurthen, M., and König, P. (2013). Where's the action? The pragmatic turn in cognitive science. *Trends Cogn. Sci.* 17, 202–209. doi: 10.1016/j.tics.2013.03.006

Engström, J., Markkula, G., Victor, T., and Merat, N. (2017). Effects of cognitive load on driving performance: the cognitive control hypothesis. *Hum. Factors* 59, 734–764. doi: 10.1177/0018720817690639

Fallgatter, A. J., Aranda, D. R., Bartsch, A. J., and Herrmann, M. J. (2002). Long-term reliability of electrophysiologic response control parameters. *J. Clin. Neurophysiol.* 19, 61–66. doi: 10.1097/00004691-200201000-00008

Fernández, T., Harmony, T., Rodríguez, M., Reyes, A., Marosi, E., and Bernal, J. (1993). Test-retest reliability of EEG spectral parameters during cognitive tasks: I absolute and relative power. *Int. J. Neurosci.* 68, 255–261. doi: 10.3109/00207459308994280

Gasser, T., Bächer, P., and Steinberg, H. (1985). Test-retest reliability of spectral parameters of the EEG. *Electroencephalogr. Clin. Neurophysiol.* 60, 312–319. doi: 10.1016/0013-4694(85)90005-7

Getzmann, S., Arnau, S., Karthaus, M., Reiser, J. E., and Wascher, E. (2018). Age-related differences in pro-active driving behavior revealed by EEG measures. *Front. Hum. Neurosci.* 12, 1–10. doi: 10.3389/fnhum.2018.00321

Gevins, A., Smith, M. E., McEvoy, L., and Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cereb. Cortex* 7, 374–385. doi: 10.1093/cercor/7.4.374

Hanslmayr, S., Staudigl, T., and Fellner, M.-C. (2012). Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Front. Hum. Neurosci*. 3:74. doi: 10.3389/fnhum.2012.00074

Harms, L. (1986). "Drivers' attentional responses to environmental variations: A dual-task real traffic study," in *Vision in Vehicles*, eds A. G. Gale, M. H. Freeman, C. M. Haslegrave, P. Smith, and S. P. Taylor (Amsterdam: North Holland), 131–138.

Hatz, F., Hardmeier, M., Bousleiman, H., Rüegg, S., Schindler, C., and Fuhr, P. (2015). Reliability of fully automated versus visually controlled pre- and post-processing of resting-state EEG. *Clin. Neurophysiol.* 126, 268–274. doi: 10.1016/j.clinph.2014.05.014

Herrmann, C. S., and Knight, R. T. (2001). Mechanisms of human attention: event-related potentials and oscillations. *Neurosci. Biobehav. Rev.* 25, 465–476. doi: 10.1016/S0149-7634(01)00027-6

Islam, M. R., Barua, S., Ahmed, M. U., Begum, S., Aric,ò, P., Borghini, G., et al. (2020). A novel mutual information based feature set for drivers' mental workload evaluation using machine learning. *Brain Sci.* 10:551. doi: 10.3390/brainsci10080551

Jensen, O., and Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *Eur. J. Neurosci.* 15, 1395–1399. doi: 10.1046/j.1460-9568.2002.01975.x

Karthaus, M., Wascher, E., and Getzmann, S. (2018). Proactive vs. reactive car driving: EEG evidence for different driving strategies of older drivers. *PLoS ONE* 13, e0191500–e0191512. doi: 10.1371/journal.pone.0191500

Kilner, J. M., Mattout, J., Henson, R., and Friston, K. J. (2005). Hemodynamic correlates of EEG: a heuristic. *Neuroimage* 28, 280–286. doi: 10.1016/j.neuroimage.2005.06.008

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Brain Res. Rev.* 29, 169–195. doi: 10.1016/S0165-0173(98)00056-3

Labounek, R., Bridwell, D. A., Mareček, R., Lamoš, M., Mikl, M., Brázdil, M., et al. (2019). "Stable EEG spatiospectral sources using relative power as group-ICA input," in *World Congress on Medical Physics and Biomedical Engineering 2018*, eds L. Lhotsha, L. Sukupova, I. Lackovic, and G. S. Ibbott (Singapore: Springer), 125–128.

Labounek, R., Lamo,š, M., Mareček, R., Brázdil, M., and Jan, J. (2015). Exploring task-related variability in fMRI data using fluctuations in power spectrum of simultaneously acquired EEG. *J. Neurosci. Methods* 245, 125–136. doi: 10.1016/j.jneumeth.2015.02.016

Lal, S. K. L., and Craig, A. (2001). A critical review of the psychophysiology of driver fatigue. *Biol. Psychol.* 55, 173–194. doi: 10.1016/S0301-0511(00)00085-5

Lal, S. K. L., and Craig, A. (2005). Reproducibility of the spectral components of the electroencephalogram during driver fatigue. *Int. J. Psychophysiol.* 55, 137–143. doi: 10.1016/j.ijpsycho.2004.07.001

Lin, C. T., Chuang, C. H., Kerick, S., Mullen, T., Jung, T. P., Ko, L. W., et al. (2016). Mind-wandering tends to occur under low perceptual demands during driving. *Sci. Rep.* 6, 1–11. doi: 10.1038/srep21353

Lohani, M., Payne, B. R., and Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Front. Hum. Neurosci.* 13:57. doi: 10.3389/fnhum.2019.00057

Lopes da Silva, F. (2013). Eeg and meg: relevance to neuroscience. *Neuron* 80, 1112–1128. doi: 10.1016/j.neuron.2013.10.017

Martens, M. H., and Fox, M. R. J. (2007). Do familiarity and expectations change perception? Drivers' glances and response to changes. *Trans. Res. Part F Traffic Psychol. Behav.* 10, 476–492. doi: 10.1016/j.trf.2007.05.003

McEvoy, L. K., Smith, M. E., and Gevins, A. (2000). Test–retest reliability of cognitive EEG. *Clin. Neurophysiol.* 111, 457–463. doi: 10.1016/S1388-2457(99)00258-8

Mikkelsen, K. B., Ebajemito, J. K., Bonmati-Carrion, M. A., Santhi, N., Revell, V. L., Atzori, G., et al. (2019). Machine-learning-derived sleep–wake staging from around-the-ear electroencephalogram outperforms manual scoring and actigraphy. *J. Sleep Res.* 28:e12786. doi: 10.1111/jsr.12786

Mirkovic, B., Bleichner, M. G., De Vos, M., and Debener, S. (2016). Target speaker detection with concealed EEG around the ear. *Front. Neurosci.* 10:349. doi: 10.3389/fnins.2016.00349

Mordkoff, J. T. (2019). A simple method for removing bias from a popular measure of standardized effect size: adjusted partial eta squared. *Adv. Methods Pract. Psychol. Sci.* 2, 228–232. doi: 10.1177/2515245919855053

Mullen, T., Kothe, C., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., et al. (2014). "Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG," in *Presented at the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE*, 2184–2187.

Mullen, T. R., Kothe, C. A. E., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., et al. (2017). Real-time neuroimaging and cognitive monitoring using wearable dry EEG. *IEEE Trans. Biomed. Eng.* 62, 2553–2567. doi: 10.1109/TBME.2015.2481482

Näpflin, M., Wildi, M., and Sarnthein, J. (2008). Test–retest reliability of EEG spectra during a working memory task. *Neuroimage* 43, 687–693. doi: 10.1016/j.neuroimage.2008.08.028

Nishida, K., Razavi, N., Jann, K., Yoshimura, M., Dierks, T., Kinoshita, T., et al. (2015). Integrating different aspects of resting brain activity: a review of electroencephalographic signatures in resting state networks derived from functional magnetic resonance imaging. *Neuropsychobiology* 71, 6–16. doi: 10.1159/000363342

Onton, J., Delorme, A., and Makeig, S. (2005). Frontal midline EEG dynamics during working memory. *Neuroimage* 27, 341–356. doi: 10.1016/j.neuroimage.2005.04.014

Pacharra, M., Debener, S., and Wascher, E. (2017). Concealed around-the-ear EEG captures cognitive processing in a visual Simon task. *Front. Hum. Neurosci.* 11:290. doi: 10.3389/fnhum.2017.00290

Parada, F. J. (2018). Understanding natural cognition in everyday settings: 3 pressing challenges. *Front. Hum. Neurosci.* 12:386. doi: 10.3389/fnhum.2018.00386

Parada, F. J., and Rossi, A. (2020). Perfect timing: Mobile brain/body imaging scaffolds the 4E-cognition research program. *Eur. J. Neurosci.* doi: 10.1111/ejn.14783

Pattyn, N., Neyt, X., Henderickx, D., and Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: boredom or cognitive fatigue? *Physiol. Behav.* 93, 369–378. doi: 10.1016/j.physbeh.2007.09.016

Pauzié, A. (2008). A method to assess the driver mental workload: the driving activity load index (DALI). *IET Intelligent Trans. Syst.* 2, 315–319. doi: 10.1049/iet-its:20080023

Plechawska-Wojcik, M., Kaczorowska, M., and Zapała, D. (2019). "The artifact subspace reconstruction (ASR) for EEG signal correction. A comparative study," in: *Proceedings of 39th International Conference on Information Systems Architecture and Technology – ISAT 2018. Advances in Intelligent Systems and Computing*, eds L. Borzemski, J Swiatek, Z Wilimowska. Vol. 853 (Cham: Springer).

Puma, S., Matton, N., Paubel, P. V., Raufaste, É., and El-Yagoubi, R. (2018). Using theta and alpha band power to assess cognitive workload in multitasking environments. *Int. J. Psychophysiol.* 123, 111–120. doi: 10.1016/j.ijpsycho.2017.10.004

Rahman, N. I. A., Dawal, S. Z. M., and Yusoff, N. (2017). Subjective responses of mental workload during real time driving: a pilot field study. *IOP Conf. Series Mater. Sci. Eng.* 210:012076.

Reiser, J. E., Wascher, E., and Arnau, S. (2019). Recording mobile EEG in an outdoor environment reveals cognitive-motor interference dependent on movement complexity. *Sci. Rep.* 9, 1–14. doi: 10.1038/s41598-019-49503-4

Reiser, J. E., Wascher, E., Rinkenauer, G., and Arnau, S. (2020). Cognitive-motor interference in the wild: Assessing the effects of movement complexity on task switching using mobile EEG. *Eur. J. Neurosci.* 1–21. doi: 10.1111/ejn.14959

Rogers, J. M., Johnstone, S. J., Aminov, A., Donnelly, J., and Wilson, P. H. (2016). Test-retest reliability of a single-channel, wireless EEG system. *Int. J. Psychophysiol.* 106, 87–96. doi: 10.1016/j.ijpsycho.2016.06.006

Sammer, G., Blecker, C., Gebhardt, H., Bischoff, M., Stark, R., Morgen, K., et al. (2007). Relationship between regional hemodynamic activity and simultaneously recorded EEG-theta associated with mental arithmetic-induced workload. *Hum. Brain Mapp.* 28, 793–803. doi: 10.1002/hbm.20309

Scholz, A., Ghadiri, A., Singh, U., Wendsche, J., Peters, T., and Schneider, S. (2018). Functional work breaks in a high-demanding work environment: an experimental field study. *Ergonomics* 61, 255–264. doi: 10.1080/00140139.2017.1349938

Sterr, A., Ebajemito, J. K., Mikkelsen, K. B., Bonmati-Carrion, M. A., Santhi, N., Della Monica, C., et al. (2018). Sleep EEG derived from behind-the-ear Electrodes (cEEGrid) compared to standard polysomnography: a proof of concept study. *Front. Hum. Neurosci.* 12:452. doi: 10.3389/fnhum.2018.00452

Symeonidou, E.-R., Nordin, A., Hairston, W., and Ferris, D. (2018). Effects of cable sway, electrode surface area, and electrode mass on electroencephalography signal quality during motion. *Sensors* 18, 1–13. doi: 10.3390/s18041073

Trick, L. M., Toxopeus, R., and Wilson, D. (2010). The effects of visibility conditions, traffic density, and navigational challenge on speed compensation and driving performance in older adults. *Accident Anal. Prevent.* 42, 1661–1671. doi: 10.1016/j.aap.2010.04.005

Van Albada, S. J., Kerr, C. C., Chiang, A. K. I., Rennie, C. J., and Robinson, P. A. (2010). Neurophysiological changes with age probed by inverse modeling of EEG spectra. *Clin. Neurophysiol.* 121, 21–38. doi: 10.1016/j.clinph.2009.09.021

Van Albada, S. J., Rennie, C. J., and Robinson, P. A. (2007). Variability of model-free and model-based quantitative measures of EEG. *J. Integr. Neurosci.* 6, 279–307. doi: 10.1142/S0219635207001520

Verwey, W. B., and Veltman, H. A. (1996). Detecting short periods of elevated workload: a comparison of nine workload assessment techniques. *J. Exp. Psychol. Appl.* 2, 270–285. doi: 10.1037/1076-898X.2.3.270

Wang, Y. K., Jung, T. P., and Lin, C. T. (2018). Theta and alpha oscillations in attentional interaction during distracted driving. *Front. Behav. Neurosci.* 12:3. doi: 10.3389/fnbeh.2018.00003

Wascher, E., Arnau, S., Reiser, J. E., Rudinger, G., Karthaus, M., and Getzmann, S. (2019). Evaluating mental load during realistic driving simulations by means of round the ear electrodes. *Front. Neurosci.* 13:940. doi: 10.3389/fnins.2019.00940

Wascher, E., Getzmann, S., and Karthaus, M. (2016). Driver state examination—Treading new paths. *Accident Anal. Prevent.* 91, 157–165. doi: 10.1016/j.aap.2016.02.029

Wascher, E., Rasch, B., Sänger, J., Hoffmann, S., Schneider, D., Rinkenauer, G., et al. (2014). Frontal Theta activity reflects distinct aspects of mental fatigue. *Biol. Psychol.* 96, 57–65. doi: 10.1016/j.biopsycho.2013.11.010

Wascher, E., Reiser, J., Rinkenauer, G., Larra, M., Dreger, F., Schneider, D., et al. (2021). Neuroergonomics on the go. An evaluation of the potential of mobile EEG for work-place asessment and design. *Hum. Factors*. doi: 10.1177/00187208211007707. [Epub ahead of print].

Wilson, G. F., and Hankins, T. (1994). Eeg and subjective measures of private pilot workload. *Proc. Hum. Factors Ergonom. Soc. Ann. Meet.* 38, 1322–1325. doi: 10.1177/154193129403801916

Yamada, F. (1998). Frontal midline theta rhythm and eyeblinking activity during a VDT task and a video game: useful tools for psychophysiology in ergonomics. *Ergonomics* 41, 678–688. doi: 10.1080/001401398186847

Yanko, M. R., and Spalek, T. M. (2013). Route familiarity breeds inattention: a driving simulator study. *Accident Anal. Prevent.* 57, 80–86. doi: 10.1016/j.aap.2013.04.003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroergonomics

PERSPECTIVE
published: 28 September 2021
doi: 10.3389/fnrgo.2021.718699

Inattention and Uncertainty in the Predictive Brain

Tuomo Kujala[1]* *and Otto Lappi*[2]

[1] Cognitive Science, Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland, [2] Cognitive Science, Traffic Research Unit, Faculty of Arts, University of Helsinki, Helsinki, Finland

Negative effects of inattention on task performance can be seen in many contexts of society and human behavior, such as traffic, work, and sports. In traffic, inattention is one of the most frequently cited causal factors in accidents. In order to identify inattention and mitigate its negative effects, there is a need for quantifying attentional demands of dynamic tasks, with a credible basis in cognitive modeling and neuroscience. Recent developments in cognitive science have led to theories of cognition suggesting that brains are an advanced prediction engine. The function of this prediction engine is to support perception and action by continuously matching incoming sensory input with top-down predictions of the input, generated by hierarchical models of the statistical regularities and causal relationships in the world. Based on the capacity of this predictive processing framework to explain various mental phenomena and neural data, we suggest it also provides a plausible theoretical and neural basis for modeling attentional demand and attentional capacity "in the wild" in terms of uncertainty and prediction error. We outline a predictive processing approach to the study of attentional demand and inattention in driving, based on neurologically-inspired theories of uncertainty processing and experimental research combining brain imaging, visual occlusion and computational modeling. A proper understanding of uncertainty processing would enable comparison of driver's uncertainty to a normative level of appropriate uncertainty, and thereby improve definition and detection of inattentive driving. This is the necessary first step toward applications such as attention monitoring systems for conventional and semi-automated driving.

Keywords: driving, predictive processing, occlusion, computational modeling, appropriate uncertainty

OPEN ACCESS

Edited by:
Giovanni Vecchiato,
National Research Council of Italy, Italy

Reviewed by:
Sophie Lemonnier,
EA7312 Laboratoire de Psychologie Ergonomique et Sociale pour l'Expérience Utilisateurs (PERSEUS), France
Christophe Jallais,
Université Gustave Eiffel, France

***Correspondence:**
Tuomo Kujala
tuomo.kujala@jyu.fi

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: *01 June 2021*
Accepted: *02 September 2021*
Published: *28 September 2021*

Citation:
Kujala T and Lappi O (2021) Inattention and Uncertainty in the Predictive Brain. Front. Neuroergon. 2:718699. doi: 10.3389/fnrgo.2021.718699

INTRODUCTION

"*The output of the system is easily measured, and easily understood, but it is extremely difficult to specify what the input is that results in the observed output.*"

– Senders et al., 1967

Appropriate allocation of attention is needed for successful performance in many contexts—work, traffic, education, and sports, among others. In traffic, driver distraction is considered as a contributing factor in many accidents (Née et al., 2019). Driver distraction is one form of inattention, referring to insufficient attention allocation to activities critical for safe driving due to diverting attention to unrelated activities (Regan et al., 2011). Inattention could be also caused by, for instance, mind wandering or fatigue (Walker and Trick, 2018). Superficially, the phenomenon seems straightforward: performance errors become more likely when attention is not allocated in accordance to task demand at the right time (Fuller, 2005; Regan et al., 2011). Look more deeply, and it's a bit more complicated than that.

First, after-the-fact explanations of accidents and errors being "caused by inattention" leave many questions unanswered. Kircher and Ahlström (2017) and Regan et al. (2011) raise the issue of hindsight bias: the driver failed to give way to a bicyclist when turning, a crash occurred, and therefore the driver was "not paying enough attention." A causal theory, in contrast, requires that one be able to independently define (and measure) if an operator is attentive, whether or not this leads to a performance failure. Only then can one predict and causally explain performance by (in)attention.

Second, being "fully attentive all the time" is not a realistic goal for most people, and most of the time not necessary to achieve a high level of safety. The crash risk of an experienced driver is extremely small (e.g., 1.38 crashes/million km on urban collector roads and 0.94 crashes/million km on rural arterial roads in USA according to Forbes et al., 2012). Even if inattention is often found to be involved in a crash, the occurrence of inattention often does not lead to a crash: the vast majority of episodes of momentary inattention on the road do not lead to accident (Victor et al., 2015). Drivers are able to adapt attention between the driving task and other tasks (e.g., operating the radio, talking on the phone; Tivesten and Dozza, 2015) or adapt the driving task (e.g., speed, following distance) according to their attention level (for review see Young et al., 2007, see also Fuller, 2005; Pekkanen et al., 2017, 2018). Kircher and Ahlström (2017) call for a definition of the minimum attentional requirements of safe driving.

To arrive at such a definition, the nature of attention in driving performance (and other similar "real-world" tasks) needs to be understood, at a theoretical level, in sufficiently precise terms. Toward this end, we outline a predictive processing approach to the study of attentional demand and inattention in driving, based on neurologically-inspired theories of uncertainty processing in the human brain and experimental research combining brain imaging, visual occlusion, and computational modeling.

ATTENTION AS MANAGEMENT OF COGNITIVE RESOURCES AND UNCERTAINTY

There is a general consensus that human information processing resources are limited. There are perceptual and structural constraints in the human information processing architecture. The field of view is limited, and gaze (overt attention) is sequentially deployed to one object or location at a time (Land, 2006). Short-term or working memory capacity is limited to a small number of items that can be kept in mind simultaneously (Cowan, 2016). There are different psychological views on how attention relates to these constraints, and if it is composed of a single serial resource or multiple parallel resources (Meyer and Kieras, 1997), but its limited capacity is not in serious dispute. We consider here inattention as a form of inappropriate allocation of this limited resource in space and time.

How much attention is appropriate, and when? How should the "amount" of attention be defined in the first place? We propose that this fundamental question can be most fruitfully approached from the point of view of the unifying theory of predictive processing (Clark, 2013, 2015; Friston, 2018). The key conceptual connection is to consider the deployment of attention as management of uncertainty (Feldman and Friston, 2010), and (in)approriate attention as (in)appropriate uncertainty. In this framework, complete certainty is an unattainable ideal, just as being "fully attentive all the time" is—but there is a rational way to optimally take into account uncertainty in observations and in internal models in one's beliefs and in one's actions. This (Bayesian inference) is the core of the predictive processing theory (Clark, 2013, 2015).

Recent developments in cognitive science have led to suggestions that human cognition is just such an advanced prediction engine (**Figures 1A,B**; see Rao and Ballard, 1999; Friston, 2005, 2009, 2010, 2018; Hohwy, 2013; Clark, 2015). The function of this prediction engine is to support perception and action by continuously matching incoming sensory input against predictions of the input generated by a hierarchy of generative internal models representing statistical and causal regularities in the world. Prediction error is used as a learning signal to update the models. The generative models evolve iteratively by feedback (i.e., prediction error). The approach is based on well-understood concepts from signal processing theory and machine learning. Internal model update is Bayesian belief update for which computationally tractable approximations are known (e.g., for linear systems, the Kalmán filter), and for which plausible neurobiological implementations have been proposed (for review see Friston, 2010, 2018).

Based on the predictive processing theory (Clark, 2015), we will assume that the key cognitive functions of attention in this framework are to:

1. control memory recall, that is, the generation of top-down predictions to match against perceptual feedback, and
2. direct the active sampling of perceptual information, that is, bottom-up prediction error that reduces uncertainty about the situation (e.g., through eye-movements).

Attention may also affect, for example, how internal models are updated (given that this requires cognitive resources—an engaged driver will learn more and faster), and have many other functions we do not consider further here.

The core idea of predictive processing is that the brain takes into account the uncertainties of its own models and the incoming sensory information, and tries to strike an optimal balance between these two sources of information, the top-down and the bottom-up. Here, we propose that attention can be understood in terms of this balancing process and inattention as inappropriate balance, from some normative perspective, such as over- or underconfidence in one's predictions in relation to environmental volatility in traffic. Accordingly, appropriate attention can be recast as reflecting appropriate uncertainty about the situation and its potential outcomes[1].

[1]This concept is similar to Lee and See (2004) "appropriate trust' or "appropriate reliance." However, uncertainty is more precise to define, and development of uncertainty easier to model computationally than trust, which is a phenomenological, much more subjective construct.

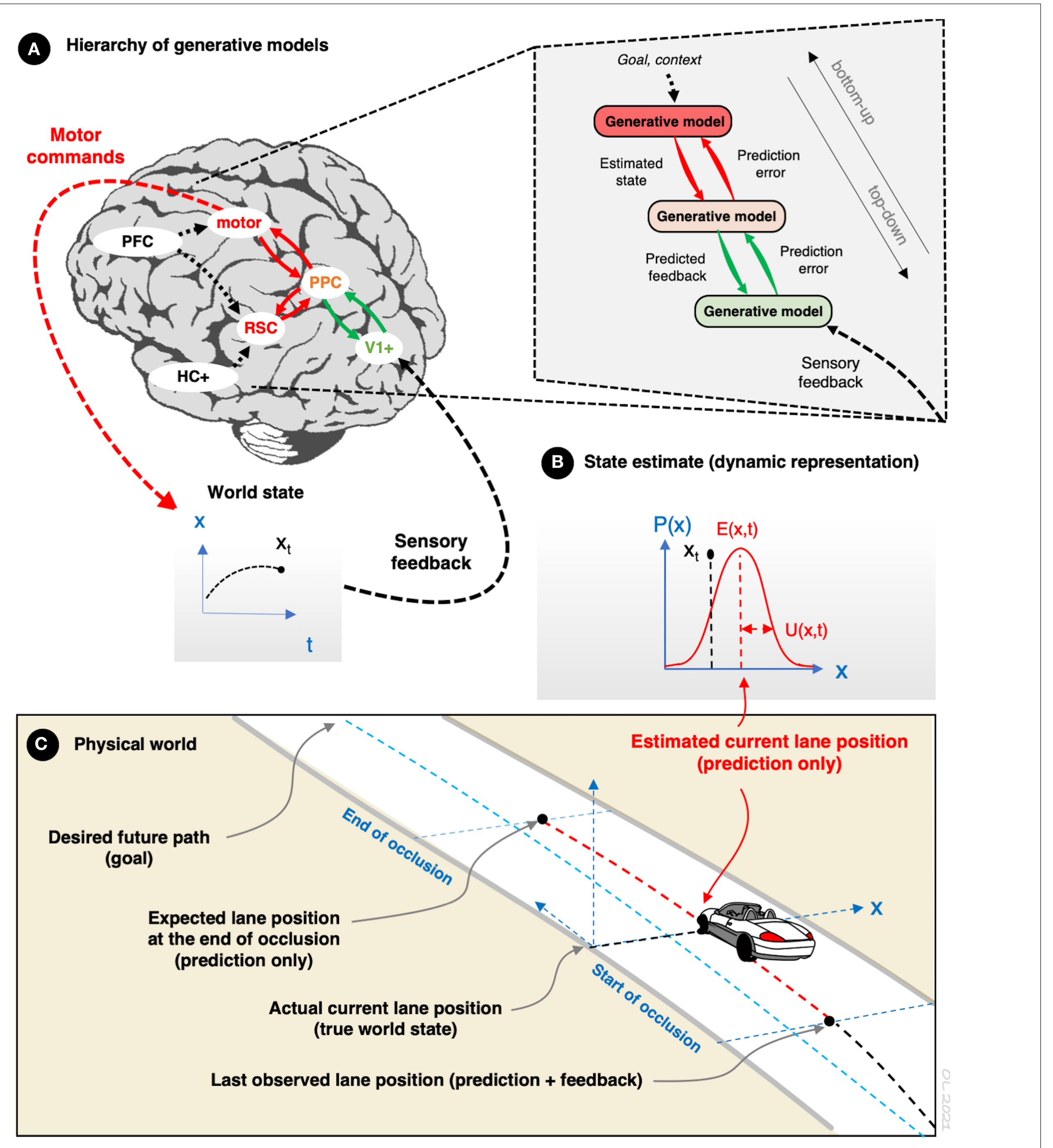


FIGURE 1 | The predictive processing framework*. **(A)** Predictions are generated by a hierarchy of generative models. Information from sensory feedback is propagated bottom-up through the hierarchy by predictive coding, and learning is based on prediction error. At each level the internal models are trying to predict their own input (only), based on memory of past events and top-down context. At the bottom (sensory) level the predictions are directly about sensor observations. The a priori prediction is compared against feedback-updated estimates, and prediction error (only) is passed forward to the higher level. The progressively higher levels behave similarly, but more abstract features of the situation are predicted (complex perceptual features, objects, events, action outcomes). At each level, prediction error is used for learning to update the internal models to determine the a priori prediction at the next time step and for similar situations in the future. Crucially, the generative models and observations are always uncertain, but the system is assumed to know this and adapt to the uncertainty in an optimal (rational) way. **(B)** Variable x_t represents a world state x at time t, which is predicted, for example lateral road position in driving. At time t, a prediction of state x can be illustrated as a probability density function, where E(x,t) is the expected value of x and U(x,t) is a dispersion measure reflecting uncertainty of the expectation, such as variance. Note that the function does not have to be Gaussian or symmetric. **(C)** Illustration of car driving on a curved road under intermittent occlusion. While occluded (red line, e.g., during off-road glances, blinks, or saccades) the estimate of the state x is updated by top-down prediction only. Artificial occlusion methods allow the study of these

(Continued)

FIGURE 1 | predictions and the associated uncertainty estimates under controlled conditions. *Driving task relevant brain regions and functions in **(A)** (Navarro et al., 2018): PFC: prefrontal cortex (goals and task context; monitoring of task performance; also representation of uncertainty and connections to limbic reward system). Motor: pyramidal and extrapyramidal motor systems, premotor and supplementary motor cortex (coordination of motor responses). HC+: hippocampus and related structures like entorhinal cortex and parahippocampal gyrus (spatial context; also memory encoding and retrieval). PPC: (posterior) parietal cortex, precuneus (multisensory spatial attention; eye movements). V1+: visual cortex and associated subcortical regions (visual perception; visual scene analysis in parahippocampal and occipital place areas). RSC: retrosplenial cortex (connecting spatial, visual and multisensory attention systems).

Our empirical approach to appropriate uncertainty builds on seminal work by Senders (1964), Senders et al. (1967), Sheridan's work (Sheridan, 1970) on supervisor's optimal sampling models and paradigms based on the visual occlusion technique (see esp. Kujala et al., 2016; Pekkanen et al., 2017). In visual occlusion experiments, the driver's sight is intermittently blocked by an occlusion visor, opaque glasses or screen on the windshield, or simply by blanking a driving simulator display (see **Figure 1C**). The driver can request a visual sample by pressing a button. Occlusion time and/or distance are calculated as the driver's estimate of spare visual capacity in driving (Safford, 1971). There is a lot of data on gross effects of various factors, such as road environment, road curvature, traffic, manoeuvre, age, and driving experience on spare visual capacity in driving (for review see Kujala et al., 2021) but a lack of a detailed understanding of the mechanisms behind these effects.

Occlusion scenarios admittedly lack some ecological face validity (information sampling in real driving is not through all or nothing occlusions), and perhaps for this reason have been less used in driver attention research than eye tracking (Kujala et al., 2021). But from the point of view of attentional processes—and especially computational modeling—the benefit is that it is not necessary to know how much and what kind of information is perceived and processed from the visual periphery (Pekkanen et al., 2017, 2018; Kircher et al., 2019; Kujala et al., 2021). Self-paced occlusion methods, in combination with other methods, allow more direct study of the predictions and the associated uncertainty estimates of the brain in controlled conditions. Note also that in natural driving, brief anticipated "occlusions" of a up to hundreds of milliseconds do occur up to several times a second (saccades, eye blinks, Land, 2006). Further, occlusion could be seen as mimicking multitasking while driving. The difference between true multitasking and occlusion on a single task is in that one can still fully focus (mentally) on the single task while occluded. Of interest could be to study the effects of additional tasks on the mental processes required for appropriate allocation of attention to the occluded task (Kujala et al., 2021).

APPROPRIATE UNCERTAINTY IN PREDICTIVE PROCESSING

Within the current framework, definition of appropriate uncertainty can be approached from at least three perspectives, each illustrated in **Figure 2**. In the example driving task (see **Figures 1C**, **2**), state x is car's lateral position. The driver has two goals: (1) D(x) is the steering goal (i.e., desired path of the car) and (2) to keep the uncertainty of lateral position U(x) under a preferred constant is the sampling "goal." D(x) also includes the implicit goal to stay on the road by remaining between road edges. A road edge defines here a task-critical threshold T(x) and (partly) a critical safety margin for the driver in the task.

Figure 2A shows an imaginary example of how, at the end of an occlusion, the driver's brain updates a prior prediction distribution about state x (car's lateral position) to a posterior distribution, based on observed feedback. It is assumed that this update is based on Approximately Bayesian Computation (the exact Bayesian distributions being intractable), that can be modeled with existing techniques such as particle filters. Note that for this reason estimation of the exact probability (or risk) of a very rare adverse event at observation, that is, x exceeding some task-critical threshold T(x) (e.g., road edge), can be highly unreliable for the brain [**Figure 2A**, (0)].

Figure 2B shows the dynamic development of expectancy, expected error, uncertainty accumulation and safety margin depletion, during an occlusion (cf. **Figure 1C**), that is, during a time interval while the driver is not observing state x. During the occlusion, the brain generates predictions, that is, samples hypotheses from the generative models, maintaining a dynamic prediction distribution about the development of state x in time.

We assume that the more hypotheses (models) sampled, the more attentive the driver is (cf. attention as control of memory recall). Paradoxically, this can mean that the more attentive, the faster the driver becomes uncertain of the development of state x during occlusion. For example, suppose the driver wishes to maintain occlusion until it is "possible" that a critical safety margin is breached [diffusion to a barrier; **Figure 2B,** (2)]. The more hypotheses, the faster the dispersion rate of the extreme values in the distribution, and therefore, the sooner the possible safety margin depletion under occlusion. This means that a more "attentive" driver will sample more frequently.

Here, it is important to notice that the expected prediction error [i.e., desired D(x) – predicted E(x,t)] can decrease in time (e.g., due to steering toward desired position) during the whole occlusion, but still uncertainty (e.g., the difference in predicted x between the most extreme hypotheses) will increase. Again, suppose the driver wishes to sample when some critical dispersion is "possibly" reached [**Figure 2B**, (1)]: the attentive driver will sample more frequently even if the driver "expects" the error at the end of occlusion to be small.

Furthermore, based on observed feedback (prediction error), attentional control of top-down processes (sampling the generative models) should adapt the number of hypotheses and their dispersion rate to be appropriately "calibrated" to the volatility of the situation, for future occlusions in similar situations. That is, a "big surprise" at the end of occlusion should

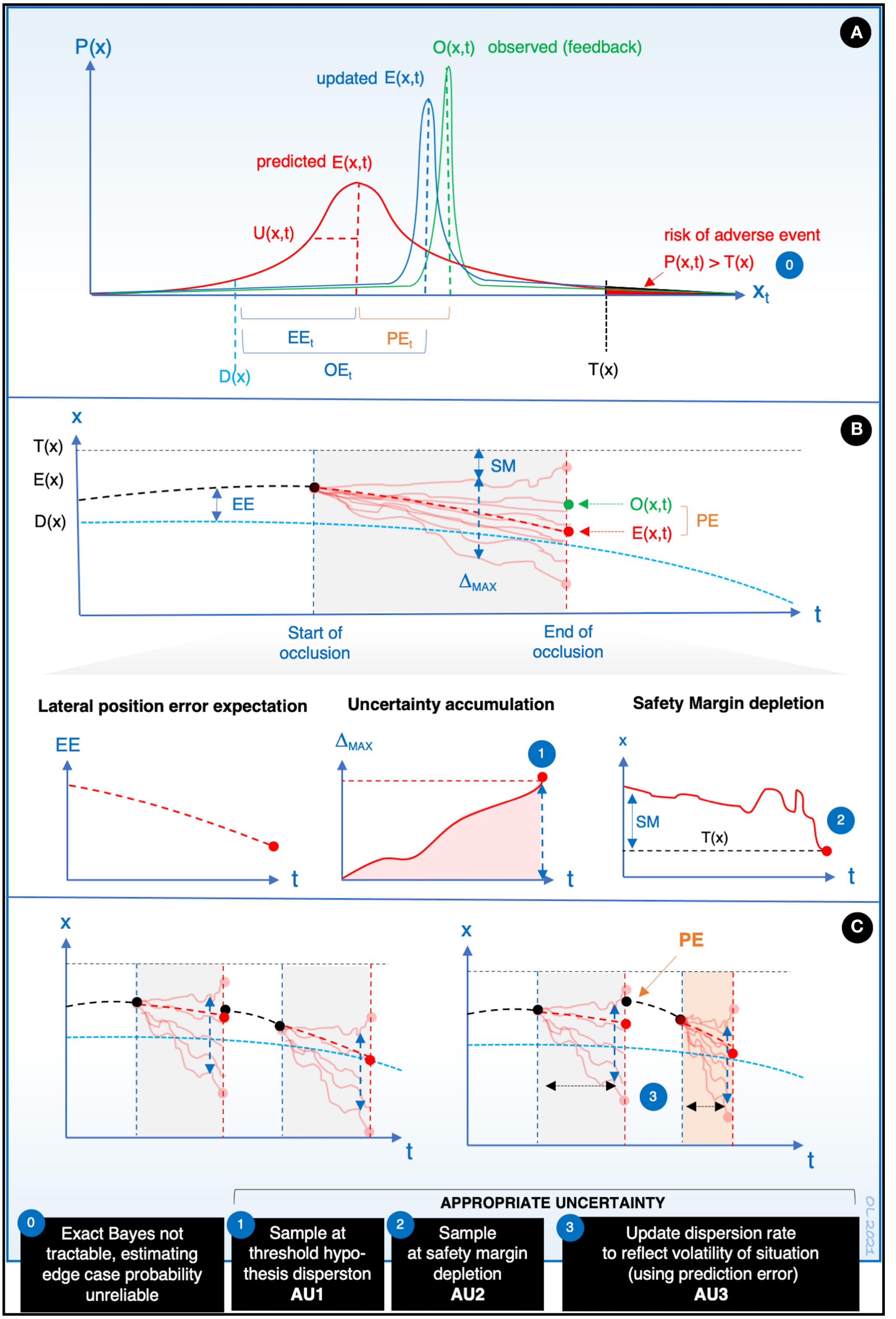


FIGURE 2 | (Approximately) Bayesian inference in occluded driving, and three ways to understand appropriate uncertainty in the predictive processing framework (AU1-3). **(A)** At the end of occlusion (t, see **Figure 1C**), state information about x is updated by combining the internal prediction (red probability density function) and

(Continued)

FIGURE 2 | observed feedback (green) to yield the posterior (blue). E(x,t) is the expected value of the prediction, "expectancy," and we refer to its difference from O(x,t), PE, as prediction error. D(x), desired x at time t (goal state); EE_t, expected error at time t (difference of expectancy and goal state); OE_t, observed error at time t. T(x), threshold × for subjective "failure" at time t (e.g., lane position where a wheel crosses a road edge), U(x,t) is a dispersion measure of the prediction distribution (here, standard deviation). Note that the functions do not have to be Gaussian or symmetric. **(B)** Illustration of the dynamic development of expectancy, expected error, uncertainty accumulation, and safety margin depletion, during an occlusion (see **Figure 1C**). This is conceived as trajectories of particles representing different hypotheses about the state that can be considered as a sample from the prediction distribution. Note that at the beginning of the occlusion, particle trajectories begin to diverge, corresponding to an increase of dispersion in the prediction distribution. Δ_{MAX} = difference in predicted x between the most extreme hypotheses (the most extreme "subjectively possible" values of x). SM, safety margin, i.e., the distance from T(x) of the most extreme hypothesis. The occlusion ends (the driver requests a sample) when some criterion is reached, such as some value of prediction distribution dispersion or the depletion of safety margin. (Note that a sample will be requested even though the expectancy approaches the goal, i.e., expected error is reduced during the occlusion). **(C)** (Appropriate) uncertainty adjustment: observation of a higher PE at the end of the previous occlusion (right) leads to higher dispersion rate in the following occlusion, and hence more frequent sampling (AU3). This is an adaptive response to situational volatility signalled by PE.

lead to more frequent visual sampling [**Figure 2C**, (3)]. This is yet another form of appropriate uncertainty. Higher volatility (as signaled by prediction error) means more unpredictable behaviour of the predicted state due to, for instance, increase in speed, variable curvature or reduced friction on the road. Higher volatility should increase uncertainty of the associated predictions (i.e., higher number of possible hypotheses in our approach). If the driver is not reactive to increased prediction error (i.e., does not adjust the uncertainty appropriately), this could lead to overconfidence in predictions, actions based on highly inaccurate state estimates and overlong occlusions, with possible negative consequences for task performance.

Figure 2 illustrates these three approaches for defining appropriate uncertainty in the predictive processing framework. First, it is rational and appropriate uncertainty (AU1), to sample feedback of x at a subjective threshold of "maximum tolerated uncertainty" U(x), provided that the threshold is appropriate for the situation, and the accumulation of uncertainty itself is appropriately calibrated. Second, it is appropriate uncertainty (AU2) to sample at a personal safety margin threshold T(x), when it is merely "subjectively possible" that the threshold is breached—regardless of the expected E(x,t) or the probability of the event (which for edge cases may be too small to estimate reliably). Third, it is appropriate uncertainty (AU3) to increase the number of hypotheses sampled from the prediction distribution, and thereby increase the dispersion rate[2] of the most extreme hypotheses and the uncertainty [U(x,t)] growth rate for a following occlusion, if the prediction error is large at the end of the previous occlusion. The driver is adapting uncertainty and thereby visual sampling on the basis of the size of the prediction error, which informs about the volatility of the situation (i.e., "uncertainty in the world"). This adjustment of dispersion rate of the hypotheses works also in the other direction; with repeated low prediction error, it is appropriate to decrease the number of the hypotheses and thereby eliminate farthest hypotheses and increase the occlusion time.

[2]Here, we assume that the brain adjusts the dispersion rate by increasing sampling of hypotheses from the prediction distribution. Another possibility is that the number of hypotheses and the dispersion rate stays at the same level but the brain lowers the threshold of maximum tolerated uncertainty, i.e., samples feedback at lower dispersion.

CONCLUSIONS

We have introduced a definition of attention as appropriate uncertainty (and inattention as inappropriate uncertainty) in predictive processing, with an application to driving under conditions of intermittent visual sampling. The novelty here is the emphasis on internal uncertainty as the basis of appropriate attention (as opposed to the false ideal of "complete certainty") and the balance between uncertainty growth rate "in the world" (i.e., volatility) and in the brain.

We have identified three criteria of appropriate uncertainty; (1) sampling perceptual feedback of state x at a personal threshold of maximum tolerated uncertainty (dispersion of predictions), (2) sampling at a personal safety margin threshold (most extreme prediction), and (3) increasing the uncertainty growth rate for a following occlusion (and for similar future situations), when the sampled prediction error is large. Violation against any of these rational behaviours can be seen as inappropriate uncertainty, and inattention (or excessive attention) toward state x.

Intuitively, the idea is that the uncertainty of, for instance, a car driver, should rise at an appropriate time and it should either grow or decrease appropriately based on changes in situational factors, such as one's own speed, relative speeds, and positions and behaviours of surrounding vehicles. It is not irrational to tolerate some uncertainty (or "risk"), which is unavoidable.

This definition suggests that "being attentive" does not mean that you are constantly processing as much task-relevant information as you possibly can, but that you are processing it to a sufficient degree to succeed in the task, based on your personal goals, previous experiences and while being sensitive to changes in environmental volatility (signaled by prediction error). Attentiveness is also not only about fixating something foveally but about processing the information and making appropriate adjustments to the uncertainty of predictions. In this framework, both overconfidence (too little uncertainty) and underconfidence (too much uncertainty) are suboptimal for the performance of a human operator (cf. Engström et al., 2018).

If the brain is indeed "Bayesian," then these sorts of processes should be the core function of the brain (Clark, 2015; Friston, 2018). That is, if the predictive processing approach holds water, then handling uncertainty and prediction error characterizes operations at all levels of neural sensory and motor hierarchies. Brain imaging research on decision making under risk and uncertainty (often under the umbrella term "neuroeconomics") has begun to reveal some specific brain structures that may play a central role in the representation of uncertainty (risk, volatility). These relate especially to the monoamine systems (norepinephrine and dopamine) and limbic structures such as the amygdala and the cingulate and orbitofrontal cortices (e.g., Angela and Dayan, 2005: acetylcholine and norepinephrine signals, Doya, 2008: norepinephrine and the orbitofrontal cortex, Rushworth and Behrens, 2008: prefrontal and cingulate cortex, Payzan-LeNestour et al., 2013: multiple distinct cortical areas and the locus coeruleus, Gordon et al., 2017: signal-to-noise-ratio in semantic wavelet induced frequency tagging, SWIFT). How this research relates to the neural substrates of driving (for review see Lappi, 2015; Navarro et al., 2018) remains an open question beyond the scope of this paper. However, from the predictive processing point of view the prediction would be that the hierarchy of networks sketched in **Figure 1A** (as identified in the meta-analysis of Navarro et al., 2018) would be a hierarchy of (top-down) predictions and (bottom-up) prediction errors. There are also uncertainty-based approaches to modeling cognitive processes that are not based on the predictive processing theories (e.g., Renninger et al., 2005; Vilares and Kording, 2011; Meyniel and Dehaene, 2017) but which might be compatible with the current approach.

Our approach introduces testable assumptions, hypotheses and novel research questions. We assumed that following prediction error the brain allocates attention (i.e., cognitive capacity) during occlusion by increasing sampling of hypotheses from the prediction distribution. Alternatively, the brain could choose to sample feedback (i.e., remove occlusion) at lower dispersion (i.e., at lower uncertainty threshold). Increased number of hypotheses with decreased occlusion time should become visible in neural correlates associated with processing of the hypotheses (cf. N1: Näätänen, 1992; P3b: Polich, 2007). Experimental designs that utilize additional tasks during occlusion could reveal how the additional tasks affect the mental processes of, for instance, hypothesis generation for the occluded task, and thereby, adjustments of uncertainty. However, the question to what extent "cognitive load" from secondary tasks relies on the same cognitive capacity as the primary (driving) task is a problem that is not yet well-understood. Besides multitasking, the effects of, for instance, cumulating driving experience, mind wandering and fatigue on uncertainty adjustment ability should be studied (and modeled). The most fundamental prediction from the theoretical approach is that when a driver is appropriately attentive toward a task-relevant state x, the size of prediction error at observation of x as detected from its neural correlates (e.g., Angela and Dayan, 2005: norepinephrine signals, Payzan-LeNestour et al., 2013: the locus coeruleus) should correlate with the following change in the sampling rate of the state (e.g., glancing frequency).

We believe that this kind of approach—combining a theoretical approach based on solid modeling concepts with a plausible physiological basis with a careful and accurate measurement and analysis of ecologically representative situations—has the potential to take the study of cognition and the brain out of the laboratory, and to address "real world" problems. These include, but are not limited to, ergonomics, human performance, attention monitoring, and safety in manual and automated driving. The approach is applicable to tasks and scenarios beyond lane keeping—and driving. For instance, driver's longitudinal control in a car following task (where x = safety distance or time-to-collision) can be computationally modeled, and has actually been modeled, as management of uncertainty (Johnson et al., 2014; Pekkanen et al., 2018).

Potential future applications of the proposed research approach include driver attention monitoring systems for conventional and semi-automated driving (Lenné et al., 2020). A proper understanding of uncertainty processing in the brain could enable comparison of driver's uncertainty to a normative level of appropriate uncertainty, and thereby improve the definition and detection of inattentive driving. However, the normative criterion for appropriate uncertainty must make theoretical sense, and it has to be well-defined. The outlined approach holds promise for delivering such a definition.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

The original idea for the paper and the term appropriate uncertainty came from TK. All authors contributed equally to the conception of the paper, development of the argument, and writing of the final version.

FUNDING

The research was partly funded by Academy of Finland (Sense of Space/334192 and Appropriate Uncertainty in Manual and Automated Driving/343259).

REFERENCES

Angela, J. Y., and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692. doi: 10.1016/j.neuron.2005.04.026

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/S0140525X12000477

Clark, A. (2015). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780190217013.001.0001

Cowan, N. (2016). *Working Memory Capacity: Classic Edition*. New York, NY: Psychology Press. doi: 10.4324/9781315625560

Doya, K. (2008). Modulators of decision making. *Nat. Neurosci.* 11, 410–416. doi: 10.1038/nn2077

Engström, J., Bärgman, J., Nilsson, D., Seppelt, B., Markkula, G., Piccinini, G. B., et al. (2018). Great expectations: a predictive processing account of automobile driving. *Theor. Issues Ergon. Sci.* 19, 156–194. doi: 10.1080/1463922X.2017.1306148

Feldman, H., and Friston, K. (2010). Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* 4:215. doi: 10.3389/fnhum.2010.00215

Forbes, G., Gardner, T., McGee, H. W., and Srinivasan, R. (2012). *Methods and Practices for Setting Speed Limits: An Informational Report (No. FHWA-SA-12-004)*. Federal Highway Administration. Office of Safety.

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* 13, 293–301. doi: 10.1016/j.tics.2009.04.005

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Friston, K. (2018). Does predictive coding have a future? *Nat. Neurosci.* 21, 1019–1021. doi: 10.1038/s41593-018-0200-7

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accid. Anal. Prevent.* 37, 461–472. doi: 10.1016/j.aap.2004.11.003

Gordon, N., Koenig-Robert, R., Tsuchiya, N., Van Boxtel, J. J., and Hohwy, J. (2017). Neural markers of predictive coding under perceptual uncertainty revealed with Hierarchical Frequency Tagging. *Elife* 6:e22749. doi: 10.7554/eLife.22749.013

Hohwy, J. (2013). *The Predictive Mind*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199682737.001.0001

Johnson, L., Sullivan, B., Hayhoe, M., and Ballard, D. (2014). Predicting human visuomotor behaviour in a driving task. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20130044. doi: 10.1098/rstb.2013.0044

Kircher, K., and Ahlström, C. (2017). Minimum required attention: a human-centered approach to driver inattention. *Hum. Factors* 59, 471–484. doi: 10.1177/0018720816672756

Kircher, K., Kujala, T., and Ahlström, C. (2019). On the difference between necessary and unnecessary glances away from the forward roadway: an occlusion study on the motorway. *Hum. Factors* 62, 1117–1131. doi: 10.1177/0018720819866946

Kujala, T., Kircher, K., and Ahlström, C. (2021). A review of occlusion as a tool to assess attentional demand in driving. *Hum. Factors* doi: 10.1177/00187208211010953. [Epub ahead of print].

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

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Prog. Retin. Eye Res.* 25, 296–324. doi: 10.1016/j.preteyeres.2006.01.002

Lappi, O. (2015). The racer's brain - how domain expertise is reflected in the neural substrates of driving. *Front. Hum. Neurosci.* 9:635. doi: 10.3389/fnhum.2015.00635

Lee, J. D., and See, K. A. (2004). Trust in automation: designing for appropriate reliance. *Hum. Factors* 46, 50–80. doi: 10.1518/hfes.46.1.50.30392

Lenné, M. G., Roady, T., and Kuo, J. (2020). "Driver state monitoring for decreased fitness to drive," in *Handbook of Human Factors for Automated, Connected, and Intelligent Vehicles,* eds D. Fisher, W. J. Horrey, J. D. Lee, and M. A. Regan (Boca Raton, FL: CRC Press) 247–262. doi: 10.1201/b21974-11

Meyer, D. E., and Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: part 1. basic mechanisms. *Psychol. Rev.* 104, 3–65. doi: 10.1037/0033-295X.104.1.3

Meyniel, F., and Dehaene, S. (2017). Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proc. Nat. Acad. Sci. U.S.A.* 114, E3859–E3868. doi: 10.1073/pnas.1615773114

Näätänen, R. (1992). *Attention and Brain Function*. Hillsdale, NJ: Lawrence Erlbaum Associates; Psychology Press.

Navarro, J., Reynaud, E., and Osiurak, F. (2018). Neuroergonomics of car driving: a critical meta-analysis of neuroimaging data on the human brain behind the wheel. *Neurosci. Biobehav. Rev.* 95, 464–479. doi: 10.1016/j.neubiorev.2018.10.016

Née, M., Contrand, B., Orriols, L., Gil-Jardiné, C., Galéra, C., and Lagarde, E. (2019). Road safety and distraction, results from a responsibility case-control study among a sample of road users interviewed at the emergency room. *Accid. Anal. Prevent.* 122, 19–24. doi: 10.1016/j.aap.2018.09.032

Payzan-LeNestour, E., Dunne, S., Bossaerts, P., and O'Doherty, J. P. (2013). The neural representation of unexpected uncertainty during value-based decision making. *Neuron* 79, 191–201. doi: 10.1016/j.neuron.2013.04.037

Pekkanen, J., Lappi, O., Itkonen, T. H., and Summala, H. (2017). Task-difficulty homeostasis in car following models: experimental validation using self-paced visual occlusion. *PLoS ONE* 12:e0169704. doi: 10.1371/journal.pone.0169704

Pekkanen, J., Lappi, O., Rinkkala, P., Tuhkanen, S., Frantsi, R., and Summala, H. (2018). A computational model for driver's cognitive state, visual perception and intermittent attention in a distracted car following task. *R. Soc. Open Sci.* 5:180194. doi: 10.1098/rsos.180194

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Rao, R. P., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580

Regan, M. A., Hallet, C., and Gordon, C. P. (2011). Driver distraction and driver inattention: definition, relationship and taxonomy. *Accid. Anal. Prevent.* 43, 1771–1781. doi: 10.1016/j.aap.2011.04.008

Renninger, L. W., Coughlan, J. M., Verghese, P., and Malik, J. (2005). "An information maximization model of eye movements," in *Advances in Neural Information Processing Systems,* 1121–1128.

Rushworth, M. F., and Behrens, T. E. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.* 11, 389–397. doi: 10.1038/nn2066

Safford, R. R. (1971). *Visual Spare Capacity in Automobile Driving and Its Sensitivity to Carboxyhemoglobin*. Columbus, OH: The Ohio State University.

Senders, J. W. (1964). The human operator as a monitor and controller of multidegree of freedom systems. *IEEE Trans. Hum. Factors Electron.* 1, 2–5. doi: 10.1109/THFE.1964.231647

Senders, J. W., Kristofferson, A. B., Levison, W. H., Dietrich, C. W., and Ward, J. L. (1967). The attentional demand of automobile driving. *Highway Res. Rec.* 195, 15–33.

Sheridan, T. B. (1970). On how often the supervisor should sample. *IEEE Trans. Syst. Sci. Cybern.* 6, 140–145. doi: 10.1109/TSSC.1970.300289

Tivesten, E., and Dozza, M. (2015). Driving context influences drivers' decision to engage in visual-manual phone tasks: evidence from a naturalistic driving study. *J. Safety Res.* 53, 87–96. doi: 10.1016/j.jsr.2015.03.010

Victor, T., Dozza, M., Bärgman, J., Boda, C. N., Engström, J., Flannagan, C., et al. (2015). *Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk (No. SHRP 2 Report S2-S08A-RW-1).* Transportation Research Board. doi: 10.17226/22297

Vilares, I., and Kording, K. (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Ann. N.Y. Acad. Sci.* 1224:22. doi: 10.1111/j.1749-6632.2011.05965.x

Walker, H. E., and Trick, L. M. (2018). Mind-wandering while driving: the impact of fatigue, task length, and sustained attention abilities. *Transport. Res. Part F* 59, 81–97. doi: 10.1016/j.trf.2018.08.009

Young, K., Regan, M., and Hammer, M. (2007). "Driver distraction: a review of the literature," in *Distracted Driving,* eds I. J. Faulks, M. Regan, M. Stevenson, J. Brown, A. Porter and J. D. Irwin (Sydney, NSW: Australasian College of Road Safety), 379–405.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroergonomics

PERSPECTIVE
published: 19 November 2021
doi: 10.3389/fnrgo.2021.786674

European NCAP Program Developments to Address Driver Distraction, Drowsiness and Sudden Sickness

Rikard Fredriksson[1,2,3†], *Michael G. Lenné*[4,5*†], *Sjef van Montfort*[6] and *Colin Grover*[3,7]

[1] Swedish Transport Administration, Skövde, Sweden, [2] Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Göteborg, Sweden, [3] European New Car Assessment Programme (Euro NCAP), Leuven, Belgium, [4] Monash University Accident Research Centre, Monash University, Melbourne, VIC, Australia, [5] Seeing Machines, Canberra, ACT, Australia, [6] TNO, Integrated Vehicle Safety Department, Helmond, Netherlands, [7] Thatcham Research, Berkshire, United Kingdom

OPEN ACCESS

Edited by:
Christer Ahlström,
Swedish National Road and Transport Research Institute, Sweden

Reviewed by:
Christopher Watling,
Queensland University of Technology, Australia
Daniel M. Roberts,
Proactive Life, Inc., United States

***Correspondence:**
Michael G. Lenné
michael.lenne@monash.edu

† These authors share first authorship

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: 30 September 2021
Accepted: 28 October 2021
Published: 19 November 2021

Citation:
Fredriksson R, Lenné MG, van Montfort S and Grover C (2021) European NCAP Program Developments to Address Driver Distraction, Drowsiness and Sudden Sickness. Front. Neuroergon. 2:786674. doi: 10.3389/fnrgo.2021.786674

Driver distraction and drowsiness remain significant contributors to death and serious injury on our roads and are long standing issues in road safety strategies around the world. With developments in automotive technology, including driver monitoring, there are now more options available for automotive manufactures to mitigate risks associated with driver state. Such developments in Occupant Status Monitoring (OSM) are being incorporated into the European New Car Assessment Programme (Euro NCAP) Safety Assist protocols. The requirements for OSM technologies are discussed along two dimensions: detection difficulty and behavioral complexity. More capable solutions will be able to provide higher levels of system availability, being the proportion of time a system could provide protection to the driver, and will be able to capture a greater proportion of complex real-word driver behavior. The testing approach could initially propose testing using both a dossier of evidence provided by the Original Equipment Manufacturer (OEM) alongside selected use of track testing. More capable systems will not rely only on warning strategies but will also include intervention strategies when a driver is not attentive. The roadmap for future OSM protocol development could consider a range of known and emerging safety risks including driving while intoxicated by alcohol or drugs, cognitive distraction, and the driver engagement requirements for supervision and take-over performance with assisted and automated driving features.

Keywords: distraction, drowsiness, driver monitoring, test protocols, consumer testing, NCAP, vehicle safety, road safety

THE NEED FOR OCCUPANT STATUS MONITORING

Driver distraction and drowsiness remain significant contributors to death and serious injury on roads around the world. Recent data from Europe and Australia confirm that approximately 25% of crashes involve drowsiness, and that distraction and inattention are factors in 29–48% of fatal and serious injury crashes (Sundfør et al., 2019; Fitzharris et al., 2020; European Commission, 2021a). In 2019 in the United States nearly 4,000 fatalities (11% of the total) and over 400,000 injuries were attributed to distraction or drowsiness (NHTSA, 2020, 2021). These numbers are likely to be underestimates given the difficulty of identifying crash causation with these factors. Sudden sickness is also a common cause of fatal crashes. In around 10% of fatal crashes in Sweden, and 6% of injury crashes in Australia, the driver suddenly became severely ill and lost control of the car (Fitzharris et al., 2020; Trafikanalys, 2021).

Road crashes attributed to distraction and drowsiness are long-standing issues in road safety strategies around the world. Road safety countermeasures have educated the public to the dangers of impaired driving and improved road infrastructure and occupant protection. Today there is even greater competition for a driver's attention. Competition for attention stems from external influences such as an increasingly busy, urbanized traffic environment and roadside (dynamic) advertising, alongside the proliferation of personal mobile devices and the "always on" society.

Managing risks in real time associated with distraction and drowsiness, as is done by intelligent speed adaptation for speeding behavior for example, has historically not been technologically possible. There is much research available now that supports the use of direct monitoring approaches, such as camera-based OSM, and that has informed the development of European Commission regulations mandating this type of technology in future years (Hynd et al., 2015). Apart from the obvious use of driver monitoring cameras to detect distraction and drowsiness, indirect symptoms of sudden sickness and driving under influence (DUI) can also be captured (e.g., head falling down or drowsiness) by the same technology and create an added benefit for these areas (Lenné, 2021).

EUROPEAN NCAP ROADMAP AND OBJECTIVES

Each year the European New Car Assessment Programme (Euro NCAP) tests all new high volume selling car models (>90% of cars sold have a rating) to provide consumer information regarding the overall safety of these cars. A total star rating is based on four areas: Adult occupant, Child occupant, Vulnerable road user and Safety assist. Protocols are typically updated every 2 years to increase the safety level. Major changes to these are laid out in a roadmap every 5 years. Under the current Euro NCAP roadmap (Euro NCAP, 2017) direct driver monitoring will be required from 2023 onwards to get a full score in the Occupant Status Monitoring (OSM) area as part of the Safety Assist Protocol. Providing a warning to drivers is important, however a stronger safety benefit will be seen if OSM is integrated with ADAS such that ADAS can become more sensitive if the driver is showing signs of inattention, drowsiness or sudden sickness. It is an important complement to the already existing areas of passive and active protection and driver support in areas such as Speed Assist systems.

DIMENSIONS OF OSM CAPABILITY

Euro NCAP's objective is to provide a strong safety outcome without over trust and an acceptable user experience to support consumer acceptance. This requires thinking about the behaviors that will be captured under the Euro NCAP program and setting definitions and test scenarios that support the stated objectives. There are two key dimensions to understanding OSM capability: detection difficulty and behavioral complexity.

The ability to detect and track the driver reliably in more complex environments equates to system availability and the proportion of time a system could provide protection to the driver. A less capable technology might be able to track in constant and less challenging environmental conditions seen in a driving simulator laboratory, but performance would degrade markedly in variable and bright lighting conditions experienced regularly in on-road driving. Particular aspects of driver appearance can also challenge performance that include eye shape and skin texture along with the driver seating position (typically indicated by driver height). Increased capability on this dimension is evident by high levels of detection accuracy with a wider range of "noise factors" that include sunglasses, hats, and masks for example.

The more recent academic and industry focus has been on defining the behaviors linked to increased risk and in developing solutions to address them. The simplest and most well-understood type of distraction behavior is a single long glance away from the roadway and is associated with increased crash risk (Victor et al., 2015). However, not all distraction meets this simplistic behavioral definition. More complex distraction behaviors are evident when drivers engage in secondary tasks such as phone use while driving. Drivers often engage in visual time sharing, where attention is split between driving and a secondary task, often up to 20–30 s (Lenné et al., 2020). This concept is recognized in several published distraction models (Seppelt et al., 2017; Kircher et al., 2020), and is important to capture to maximize safety outcomes.

The movement of a driver's head and eyes is also important. For glances that are a smaller visual angle from the forward roadway drivers typically will engage in what is termed "lizard" glance behavior. Here the drivers' eyes are moving but the head is relatively still (Fridman et al., 2016). In contrast, for glances to areas that are larger visual angle from the forward roadway, regions such as the side window and passenger seat, drivers typically engage in an "owl" strategy, where the shifting of visual attention is primarily achieved by head rotation followed by the eyes. **Figure 1** illustrates lizard visual behavior while using a phone and presents both eye gaze and head pose orientation for those sequences where the driver is looking at the phone (adapted from Yang et al., 2021). The drivers head pose remains orientated on-road. Only detection using eye gaze would detect this example of phone use. Detecting cell phone distraction will be significantly improved with approaches that measure visual behavior directly through eye gaze metrics rather than relying on head pose alone or indirect measures.

Drowsiness-related behaviors can also be characterized through a similar lens of increasing complexity. Simplistic measures of drowsiness may only capture eyelid behavior or indirect measurements. PERCLOS is an example of an eyelid-based metric used to establish a drowsy state (typically over 20 mins), however its performance is modest (Sommer and Golz, 2010; Jackson et al., 2016). Individual variability in drowsiness progression and symptoms mean that systems that rely on single

FIGURE 1 | Example of a driver using of a mobile phone with a driver assistance feature. Indicative metrics are presented for eye gaze (red) and head pose (green) when a driver uses a lizard glance strategy while looking at mobile phone.

drowsiness metrics are insufficient to capture drowsiness reliably (Ingre et al., 2006; Chua et al., 2014). Multiple signs of drowsiness, including blink duration, amplitude-velocity ratio and frequency and are likely to capture more patterns of drowsiness behavior (Caffier et al., 2003; Lee et al., 2016; Liang et al., 2019). The lack of defined objective drowsiness measures presents some additional challenges to those faced in monitoring distraction.

Microsleeps are included in the protocol, where a microsleep is a momentary period of sleep where the driver is unconscious. Microsleeps have traditionally been defined through Electroencephalography (EEG), with intrusions of theta waves anywhere between 3 and 15 s (Liang et al., 2019; Hertig-Godeschalk et al., 2020). EEG defined microsleeps have been linked with driver impairment and crash risk (Boyle et al., 2008; Golz et al., 2011). Microsleep identification through EEG is currently both impractical in driving and limited by the temporal capabilities and signal noise of the technology. Increasingly, behavioral characteristics of microsleeps have been linked to physiological and performance indicators of severe drowsiness, with long eye closures being the primary visual indicator of a microsleep (Buckley et al., 2016; Mulhall et al., 2020). In its simplest form an OSM detected microsleep could be triggered by a long eye closure, with eye closures >500 ms linked to measures of driver risk (Alvaro et al., 2016; Mulhall et al., 2020). However, there are a range of behaviors such as yawning and squinting that could be mis-interpreted as drowsiness-related long eye closure events. A simplistic definition would therefore produce a higher number of false alerts and would not provide high levels of driver acceptance. More complex definitions of microsleeps, such as those that accommodate additional indicators of microsleep (e.g., head nodding) or evidence of prior drowsiness, are needed to ensure that drivers are not receiving an excessive level of false alarms and to provide the intended safety benefits.

Recognizing sudden sickness is also part of the protocol and presents a unique challenge to data collection and ecological validity. Sudden sickness can be used as an umbrella term covering a variety of conditions (e.g., diabetic shock, cardiac events, seizures, etc.), where the common result is driver incapacitation. These events are unpredictable by nature, resulting in very sparse data, and therefore there is currently no method or taxonomy to detail these categories and their related behavior. It is reasonable to assume, however, that the driver is neither performing driving tasks effectively nor responding to vehicle alerts. In the early stages of implementation it is therefore reasonable to regard sudden sickness as a period of lack of response which can be implemented as an escalation of either drowsiness or distraction which goes uncorrected.

The behavior-detection matrix differentiates performance based upon the projected level of protection to the driver (**Figure 2**). A simpler detection technology (left end of x axis) with simpler behavioral features (bottom end of y axis) will only be able to reliably perform in the bottom left corner of the matrix. A more sophisticated technology with robust behavioral features will be able to perform toward the top right of the matrix, therefore providing coverage over a much greater range of scenarios and representing a superior solution. This matrix provides the basis of the range of noise variables and behaviors that are covered in the proposed protocol.

TEST METHODOLOGY

It is important that the protocol finds a balance that provides a safe and acceptable outcome for the community while implementing processes that are manageable by OEMs. It should incentivize widespread adoption while still affording opportunities for differentiation. It should encourage the implementation of systems that are not simply pure warning systems to the driver, but go further to integrate the OSM signal with other ADAS systems. Making ADAS systems such as automatic emergency braking or lane keep systems more sensitive when a driver is distracted, for example, is expected to provide both a greater safety benefit and more acceptable driver experience.

Protocol development considers a range of driver appearance and noise factors to ensure acceptable levels of system availability and thus system performance. The approach here is to test systems across the extremes of driver appearance, for example tall through to short drivers, and drivers of different ages, very young through to very old. Collecting data across these factors ensures good system availability with a wide range of seating positions and skin textures (wrinkles, baggy eyes). These appearance variables can be described very precisely, as is routinely done in published research studies, to give clear guidance to OEMs on the conditions being tested. The same philosophy holds with introducing noise factors into the testing.

Testing behaviors are the second element of the matrix presented in **Figure 2**. For distraction these behaviors include:

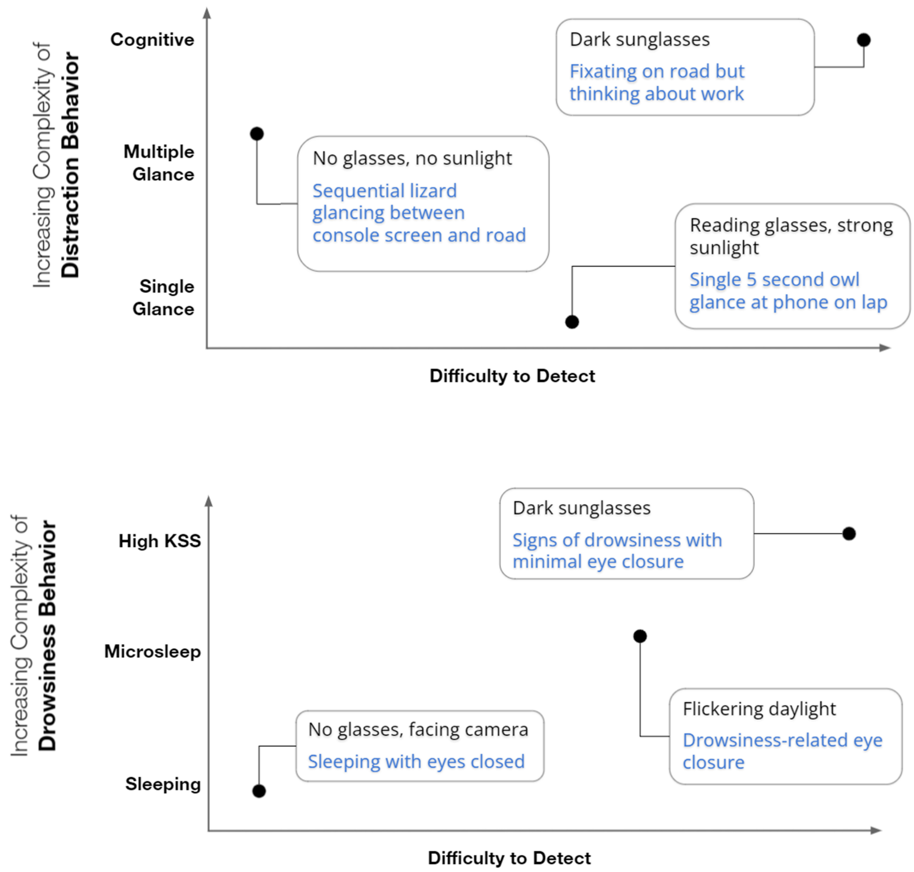


FIGURE 2 | The behavior-technology matrix for distraction (top) and drowsiness (bottom) that describes different levels of camera-based OSM technology performance. Within each figure the blue text represents the complexity of driver behavior, while the black text represents noise variables that could alter detection difficulty.

single long glances to specified driving-related and non-driving-related targets, and; visual time sharing behaviors (multiple short glances) that address risks associated with engagement in secondary activities including phone use. A test example for visual time sharing tasks could include scripted glance sequences from on-road to the console over a 10–15 s period. Testing toward the extremes of the owl and lizard glance strategies separately is a key element. This ensures that a range of individual differences in glance strategies are accommodated while also accommodating a key element that can differentiate the capability of an OSM feature. Distraction scenarios will need to be tightly prescribed and highly repeatable. Testing drowsiness-related behaviors is somewhat more complex as no single behavior or pattern is consistent across all individuals (Caffier et al., 2003; Chua et al., 2014). This makes reproducing drowsiness behaviors in a consistent manner problematic. Drowsy and microsleep data should therefore be collected from drivers that are genuinely drowsy and where this can be confirmed by validated measures [e.g., the Karolinska Sleepiness Scale (KSS) or EEG].

Ideally all testing would be conducted in test track conditions as is done with existing Euro NCAP AEB/Lane Support protocols. Track testing with a sufficient number of drivers, with different appearance, incorporating different noise factors, and testing across the range of distraction and drowsiness behaviors is not practical. The approach initially proposes testing using a dossier of evidence provided by the OEM alongside selected use of track testing. The dossier approach provides guidance to OEMs without being overly prescriptive and limiting advancements in early stage technologies, and may include recommendations of best practice guidelines for testing drowsiness, such as number of subjects and methods of inducing and validating drowsiness. Deviations from guidelines will require supporting evidence justifying the method and demonstrating comparable performance and safety benefits of the alternate approach.

Performance assessment is a key part of any testing methodology. The test philosophy of Euro NCAP is to assess how well a safety system works when needed (true positives), while the false alarm rate (false positives) is assigned to the vehicle manufacturer to address. Publicly available data for distraction algorithms estimate sensitivity performance exceeding 80% however false positives can exceed 20% (Lee et al., 2013). For drowsiness, current General Safety Regulation standards for legal acceptance are understood to place sensitivity around 40%; this mark is achievable by several algorithms but bears room for improvement at false-positive rates of 11–24% (Friedrichs and Yang, 2010; Anderson and Horne, 2013). Simply put, a vehicle with unacceptable false alarm rate will not provide an acceptable customer experience. The requirements for appropriate driver warning and vehicle intervention are directly linked to both safety outcomes and driver experience and should ensure an appropriate balance is struck between sensitivity and specificity.

FUTURE DIRECTIONS IN SENSING AND TESTING

The roadmap for future protocol development could consider a broader spectrum of behaviors and states linked to driver impairment. Alcohol and other drugs are examples given the links to fatal crashes in Europe (25% of all fatalities are alcohol-related; European Commission, 2021b), and the documented potential for real-time OSM approaches (Lenné et al., 2020; Hayley et al., 2021). We noted earlier the need for research efforts to shed new light on related features such as sudden sickness to further enhance their utility over time.

Insights from widespread implementation are likely to provide new insights for warning and intervention strategies. For drowsiness in particular combining performance and behavioral indicators, such as steering and ocular inputs for example, may improve prediction performance. From a warning and intervention viewpoint there is acknowledgment that drowsiness alerts alone will get us so far and that additional intervention strategies are needed to improve safety outcomes in the long term (Fitzharris et al., 2017). The full integration of OSM into the suite of ADAS affords an expanded range of real-time vehicle intervention strategies.

While risks associated with distraction and single long glances away from the forward roadway are well-understood, further research is needed on safety impacts of multiple glance distraction. For example: at what point during a given sequence does a driver become distracted; how is this influenced by the driver's engagement in driving and non-driving related tasks; how does the external environment influence this; and what are the links to probable crash types? Further, cognitive distraction and inattention are emerging safety issues. While reasonably well-understood in driving simulator studies, direct links to real world safety are less well-documented. Crash types here include "looked but failed to see" where a driver's visual attention can be directed on-road yet they are cognitively engaged in another activity.

Current Safety Assist protocols are designed to support drivers operating vehicles in manual driving, i.e., without assisted or automated driving functions. Driver behavior will change with increases in driver assistance and vehicle automation as drivers increasingly have the opportunity to take hands off wheel and/or eyes off road under defined conditions. It is critical to consider what safety issues these changes might introduce and how OSM can best support safe outcomes. Driver engagement is the cognitive state that is increasingly important to understand and measure here from a safety perspective (Lenné et al., 2020). Drivers need to remain sufficiently engaged and attentive to the driving task to ensure they are able to resume control should the assistance feature not perform as expected. It its simplest form, if a driver is known to be sufficiently attentive, this knowledge could be used to allow ACC to proceed from a stand-still at a red light, for example. Driver take-over readiness is key as it informs take-over performance, a safety outcome included in the planning for future Euro NCAP protocols.

There are several opportunities for researchers and industry to pursue to close some knowledge gaps. For researchers, perhaps it is about establishing the safety risks and safety scenarios for driver states that are less understood, such as cognitive distraction. Conducting in-depth crash studies to better understand the crash types and associated driver behaviors and system factors—helping to set the agenda for the problems that both technology development and safety policy should target. Continued research into the most effective warning and intervention strategies is also key. For industry there is an immediate opportunity to combine with other sensors such as child presence detection, seat belt wearing detection (advanced SBR), and occupant position and size for in-crash protection systems. There is also the opportunity to continue to push the boundaries on the safety cases that can be addressed, and the underlying technologies used to achieve this, to ensure that even greater injury reductions are realized.

CONCLUSION

Distracted and drowsy driving are highlighted as key sources of road trauma in road safety strategies around the world. These behaviors have historically been very difficult to identify when they occur while driving. OSM technologies offer new opportunities to manage driver distraction and drowsiness in real-time and thus reduce fatal and serious injury. We believe this is best achieved by combing warning and intervention strategies such as, for example, increasing the sensitivity of driver assistance systems when a driver is not attentive. The European NCAP continues to evolve its OSM protocols to recognize more advanced technologies such as driver monitoring as an integral part of upcoming rating protocols that will reward vehicle manufacturers who provide OSM features in future vehicles.

Protocols have been developed that attempt to address and mitigate the higher risk distraction and drowsiness behaviors. These protocols are likely to become effective for new vehicle models in Europe from 2023 and evolved for a 2025 update. The European NCAP roadmap in the future could include a number of known and emerging safety issues that could include cognitive distraction and take-over performance.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

RF and ML: conceptualization. ML led preparation of the original draft with RF. RF and ML: writing, reviewing, and editing. SvM and CG assisted with writing, reviewing and editing. All authors approved the final manuscript as submitted.

ACKNOWLEDGMENTS

The authors acknowledge the Euro NCAP Occupant Status Monitoring working group including ACEA, CLEPA, test laboratories and the Tier 2 supplier group for their feedback and fruitful input to the protocol currently under development. Here we also especially want to thank Richard Schram, formerly Euro NCAP for his considerable input to this work. We also acknowledge John Noble, Trey Roady, Megan Mulhall, and Grant Grubb for their contributions to the protocol rationale.

REFERENCES

Alvaro, P. K., Jackson, M. L., Berlowitz, D. J., Swann, P., and Howard, M. E. (2016). Prolonged eyelid closure episodes during sleep deprivation in professional drivers. *J. Clin. Sleep Med.* 12, 1099–1103. doi: 10.5664/jcsm.6044

Anderson, C., and Horne, J. A. (2013). Driving drowsy also worsens driver distraction. *Sleep Med.* 14, 466–468. doi: 10.1016/j.sleep.2012.11.014

Boyle, L. N., Tippin, J., Paul, A., and Rizzo, M. (2008). Driver performance in the moments surrounding a microsleep. *Transp. Res. Part F Traffic Psychol. Behav.* 11, 126–136. doi: 10.1016/j.trf.2007.08.001

Buckley, R. J., Helton, W. S., Innes, C. R. H., Dalrymple-Alford, J. C., and Jones, R. D. (2016). Attention lapses and behavioural microsleeps during tracking, psychomotor vigilance, and dual tasks. *Conscious. Cogn.* 45, 174–183. doi: 10.1016/j.concog.2016.09.002

Caffier, P. P., Erdmann, U., and Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *Eur. J. Appl. Physiol.* 89, 319–325. doi: 10.1007/s00421-003-0807-5

Chua, E. C.-P., Yeo, S.-C., Lee, I. T.-G., Tan, L.-C., Lau, P., Tan, S. S., et al. (2014). Individual differences in physiologic measures are stable across repeated exposures to total sleep deprivation. *Physiol. Rep.* 2:e12129. doi: 10.14814/phy2.12129

Euro NCAP (2017). *Euro NCAP 2025 Roadmap—In pursuit of Vision Zero*. Available online at: https://cdn.euroncap.com/media/30700/euroncap-roadmap-2025-v4.pdf

European Commission (2021a). *Road Safety Thematic Report–Fatigue. European Road Safety Observatory*. European Commission, Directorate General for Transport.

European Commission (2021b). *EC Mobility and Transport, Road Safety Website*. Available online at: https://ec.europa.eu/transport/road_safety/specialist/knowledge/alcohol_en (accessed September 28, 2021).

Fitzharris, M., Lenné, M. G., Corben, B., Arundell, T. P., Peiris, S., Lui, S., et al. (2020). *ECIS Report 1. Overview and Analysis of Crash Types, Injury Outcomes and Contributing Factors. Report # 343*. Monash University Accident Research Centre.

Fitzharris, M., Liu, S., Stephens, A. N., and 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.* 81, 71–78. doi: 10.1080/15389588.2017.1306855

Fridman, L., Lee, J., Reimer, B., and Victor, T. (2016). "Owl" and "Lizard": patterns of head pose and eye pose in driver gaze classification. *IET Comput. Vis.* 10, 308–314. doi: 10.1049/iet-cvi.2015.0296

Friedrichs, F., and Yang, B. (2010). "Drowsiness monitoring by steering and lane data based features under real driving conditions," in *Proceedings of the 18th European Signal Processing Conference* (Amsterdam), 209–213.

Golz, M., Sommer, D., Krajewski, J., Trutschel, U., and Edwards, D. (2011). Microsleep episodes and related crashes during overnight driving simulations. *Driv. Assess. Conf.* 6, 39–45. doi: 10.17077/drivingassessment.1375

Hayley, A. C., Shiferaw, B., Aitken, B., Vinckenbosch, F., Brown, T. L., and and Downey, L. A. (2021). Driver monitoring systems (DMS): the future of impaired driving management? *Traffic Inj. Prev.* 22, 313–317. doi: 10.1080/15389588.2021.1899164

Hertig-Godeschalk, A., Skorucak, J., Malafeev, A., Achermann, P., Mathis, J., and Schreier, D. R. (2020). Microsleep episodes in the borderland between wakefulness and sleep. *Sleep* 43:zsz163. doi: 10.1093/sleep/zsz163

Hynd, D., McCarthy, M., Carroll, J., Seidl, M., Edwards, M., Visvikis, C., et al. (2015). *Benefit and Feasibility of a Range of New Technologies and Unregulated Measures in the Field of Vehicle Occupant Safety and Protection of Vulnerable Road Users*. Transport Research Laboratory for the European Commission.

Ingre, M., Åkerstedt, T., Peters, B., Anund, A., and Kecklund, G. (2006). Subjective sleepiness, simulated driving performance and blink duration: examining individual differences. *J. Sleep Res.* 15, 47–53 doi: 10.1111/j.1365-2869.2006.00504.x

Jackson, M. L., Raj, S., Croft, R. J., Hayley, A. C., Downey, L. A., Kennedy, G. A., et al. (2016). Slow eyelid closure as a measure of driver drowsiness and its relationship to performance. *Traffic Inj. Prev.* 17, 251–257. doi: 10.1080/15389588.2015.1055327

Kircher, K., Kujala, T., and Ahlström, C. (2020). On the difference between necessary and unnecessary glances away from the forward roadway: an occlusion study on the motorway. *Hum. Factors* 62, 1117–1131. doi: 10.1177/0018720819866946

Lee, J. D., Moeckli, J., Brown, T. L., Roberts, S. C., Schwarz, C., Yekhshatyan, L., et al. (2013). *Distraction Detection and Mitigation Through Driver Feedback. (Report No. DOT HS 811 547A)*. National Highway Traffic Safety Administration.

Lee, M. L., Howard, M. E., Horrey, W. J., Liang, Y., Anderson, C., Shreeve, M. S., et al. (2016). High risk of near-crash driving events following night-shift work. *Proc. Nat. Acad. Sci. U.S.A.* 113:176. doi: 10.1073/pnas.1510383112

Lenné, M. (2021). *NHTSA Request for Information: Impaired Driving Technologies*. Available online at: https://www.regulations.gov/comment/NHTSA-2020-0102-0004

Lenné, M. G., Roady, T., and Kuo, J. (2020). "Driver state monitoring for decreased fitness to drive," in *Handbook of Human Factors for Automated, Connected, and Intelligent Vehicles*, eds D. Fisher, W. J. Horrey, J. D. Lee, and M. A. Regan (Boca Raton: CRC Press). doi: 10.1201/b21974-11

Liang, Y., Horrey, W. J., Howard, M. E., Lee, M. L., Anderson, C., Shreeve, M. S., et al. (2019). Prediction of drowsiness events in night shift workers during morning driving. *Accid. Anal. Prev.* 126, 105–114. doi: 10.1016/j.aap.2017.11.004

Mulhall, M. D., Cori, J., Sletten, T. L., Kuo, J., Lenné, M. G., Magee, M., et al. (2020). A pre-drive ocular assessment predicts alertness and driving impairment: a naturalistic driving study in shift workers. *Accid. Anal. Prev.* 135:105386. doi: 10.1016/j.aap.2019.105386

NHTSA (2020). *Overview of Motor Vehicle Crashes in 2019. U.S. Department of Transportation (Report DOT HS 813 060).* National Highway Traffic Safety Administration.

NHTSA (2021). *Distracted Driving 2019: Summary of Statistical Findings (Report DOT HS 813 111).* National Highway Traffic Safety Administration.

Seppelt, B., Seaman, S., Lee, J., Angell, L., Mehler, B., and Reimer, B. (2017). Glass half-full: on-road glance metrics differentiate crashes from near-crashes in the 100-Car data. *Accid. Anal. Prev.* 107, 48–62. doi: 10.1016/j.aap.2017.07.021

Sommer, D., and Golz, M. (2010). "Evaluation of PERCLOS based current fatigue monitoring technologies," in *Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology* (Buenos Aires), 4456–4459. doi: 10.1109/IEMBS.2010.5625960

Sundfør, H. B., Sagberg, F., and Høye, A. (2019). Inattention and distraction in fatal road crashes–results from in-depth crash investigations in Norway. *Accid. Anal. Prev.* 125, 152–157. doi: 10.1016/j.aap.2019.02.004

Trafikanalys (2021). *Road Traffic Injuries. (Official statistics of Sweden).* Available online at: https://www.trafa.se/en/road-traffic/road-traffic-injuries/

Victor, T., Dozza, M., Bärgman, J., Boda, C., Engström, J., Flannagan, C., et al. (2015). *Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk. SHRP 2 Report S2-S08A-RW-1.* National Academy of Sciences.

Yang, S., Shiferaw, B., Roady, T., Kuo, J., and Lenné, M. (2021). "Driver glancing like a lizard in cell phone distraction in level 2 automated driving," in *Proceedings of the 65th Internal Annual Meeting on the Human Factors and Ergonomics Society* (Baltimore).

Conflict of Interest: ML was employed by and holds shares in Seeing Machines.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroergonomics

ORIGINAL RESEARCH
published: 22 November 2021
doi: 10.3389/fnrgo.2021.682119

Risk Assessment by a Passenger of an Autonomous Vehicle Among Pedestrians: Relationship Between Subjective and Physiological Measures

Jeffery Petit[1], *Camilo Charron*[1,2] *and Franck Mars*[1*]

[1] Université de Nantes, Centrale Nantes, CNRS, LS2N, Nantes, France, [2] Université Rennes 2, Rennes, France

OPEN ACCESS

Edited by:
Lewis L. Chuang,
Ludwig Maximilian University of Munich, Germany

Reviewed by:
Klas Ihme,
Helmholtz Association of German Research Centers (HZ), Germany
Fei Yan,
University of Ulm, Germany

***Correspondence:**
Franck Mars
franck.mars@ls2n.fr

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: *17 March 2021*
Accepted: *08 October 2021*
Published: *22 November 2021*

Citation:
Petit J, Charron C and Mars F (2021) Risk Assessment by a Passenger of an Autonomous Vehicle Among Pedestrians: Relationship Between Subjective and Physiological Measures. Front. Neuroergon. 2:682119. doi: 10.3389/fnrgo.2021.682119

Autonomous navigation becomes complex when it is performed in an environment that lacks road signs and includes a variety of users, including vulnerable pedestrians. This article deals with the perception of collision risk from the viewpoint of a passenger sitting in the driver's seat who has delegated the total control of their vehicle to an autonomous system. The proposed study is based on an experiment that used a fixed-base driving simulator. The study was conducted using a group of 20 volunteer participants. Scenarios were developed to simulate avoidance manoeuvres that involved pedestrians walking at 4.5 kph and an autonomous vehicle that was otherwise driving in a straight line at 30 kph. The main objective was to compare two systems of risk perception: These included subjective risk assessments obtained with an analogue handset provided to the participants and electrodermal activity (EDA) that was measured using skin conductance sensors. The relationship between these two types of measures, which possibly relates to the two systems of risk perception, is not unequivocally described in the literature. This experiment addresses this relationship by manipulating two factors: The time-to-collision (TTC) at the initiation of a pedestrian avoidance manoeuvre and the lateral offset left between a vehicle and a pedestrian. These manipulations of vehicle dynamics made it possible to simulate different safety margins regarding pedestrians during avoidance manoeuvres. The conditional dependencies between the two systems and the manipulated factors were studied using hybrid Bayesian networks. This relationship was inferred by selecting the best Bayesian network structure based on the Bayesian information criterion. The results demonstrate that the reduction of safety margins increases risk perception according to both types of indicators. However, the increase in subjective risk is more pronounced than the physiological response. While the indicators cannot be considered redundant, data modeling suggests that the two risk perception systems are not independent.

Keywords: autonomous driving, passenger perception, risk assessment, skin conductance, driving simulator, Bayesian network

1. INTRODUCTION

Traveling in autonomous vehicles changes the driver's role when they become a passenger after ceding control to an automated system (Reilhac et al., 2016; Kyriakidis et al., 2019). Verberne et al. (2012) suggested that individuals do not change their social rules when interacting with an automated system. Basu et al. (2017) supported this idea by revealing that most drivers prefer a driving style that resembles their own. This consideration is all the more important when traveling in a dense space where different types of vulnerable road users (e.g., pedestrians and cyclists) can circulate freely. Such spaces are beginning to appear in Europe and are known as shared spaces (Hamilton-Baillie, 2008). They are designed to eliminate any segregation between road users (e.g., through a lack of signs and road markings). As a result, the fluidity of mobility in these areas essentially relies on social conventions, especially in crowded situations. One of the objectives of this urban design is to enable drivers to better integrate into multi-user environments by reducing vehicle speeds and improving traffic flow (Hamilton-Baillie, 2008; Kaparias et al., 2012; Moody and Melia, 2014). However, such designs have also introduced a new challenge for autonomous vehicles that must navigate among other users in non-signalled areas (e.g., the problem of crowd navigation as discussed by Bresson et al., 2019). The trajectories followed by the vehicle to navigate within this type of environment must remain acceptable to the users around the vehicle but also to the driver-passenger inside. Many studies currently investigate the communication between the autonomous vehicle and pedestrians through external human-machine interfaces (Faas et al., 2020; Métayer and Coeugnet, 2021), but it is also important to ensure that the passenger does not fear a collision risk (e.g., when the vehicle adopts proactive navigation, Kabtoul et al., 2020). It is therefore essential to study what will determine the acceptability of the vehicle's trajectory relative to other road users.

1.1. Vehicle Dynamics and Passenger Risk Perception

Gibson and Crooks (1938) proposed the existence of a dynamic space that the driver perceives as an area in which they can navigate safely. The authors named it the "field of safe travel." This zone represents an envelope of acceptable trajectories for a vehicle. It depends on the driver's experience, the safety distances they wish to respect and their perception of the size of the car, among other factors. Based on these considerations, Kolekar et al. (2020a,b) proposed the driving risk field to model the importance that a driver ascribes to an obstacle that blocks a straight trajectory. In their work, the authors built upon Näätänen and Summala's theory (Näätänen and Summala, 1976), who defined perceived risk as a function of both the subjective importance given to a hazard and the consequences that this hazard could pose. Kolekar et al. (2020a,b) hypothesized that the subjective importance that is given to a risk is proportional to the driver's reaction at the steering wheel when confronted with an obstacle in their trajectory. Using this perspective, the authors developed a measure proportional to the perceived risk if a hazard's consequences remain unchanged (e.g., collision with the same obstacle). Other researchers have found that the values of time-to-collision (TTC) or time headway when following a vehicle or approaching a slower obstacle highly correlate with a driver's perception of a collision (Vogel, 2003; Chen et al., 2016; Zhao et al., 2020). Researchers have particularly investigated TTC in the literature and have demonstrated that it is directly perceived through retinal expansion (Lee, 1976; Bootsma et al., 1997; Bootsma and Craig, 2003). When approaching an obstacle, an autonomous vehicle must initiate an avoidance manoeuvre to avoid a collision. When the path of the vehicle deviates from the obstacle, measures such as the TTC or time headway are no longer relevant while the vehicle continues to approach. In this case, new metrics must be used to study risk perception. Ferrier-Barbut et al. (2018) revealed the existence of a comfort zone that is perceived by the passenger of an autonomous vehicle when the vehicle is passing close to a pedestrian. This suggests that absolute distance is a factor in the passenger's perception. During an avoidance manoeuvre, this distance corresponds to the lateral distance (which is also referred to as the offset) between the vehicle and the obstacle. In summary, the study of a passenger's risk perception that involves an autonomous vehicle must integrate vehicle-environment dynamics.

1.2. The Hypothesis of Two Risk Perception Systems

Slovic et al. (2004) described two risk perception systems that are involved in evaluation and decision-making processes when an individual is faced with a potential hazard. "Risk as analysis," according to the authors, is a system of risk perception that is based on conscious reasoning and uses formal logic. This method for perceiving risk is a common conception in the scientific literature. It assumes that individuals perceive risk by estimating the product of the probability of a hazard and its consequences. However, this type of risk perception, which is slow and costly in cognitive resources, would not be solicited in the event of an imminent threat. Slovic et al. (2004) suggested the existence of a second system of perception that is predominant in this type of situation, which they named "risk as feelings." This system comprises a quick and reactive way of perceiving risk and is intuitive in nature as it is based on affects. This duality of risk perception aligns with the vision of Loewenstein et al. (2001), who suggested that decision-making results not only from cognitive processing but also from an instinctive and spontaneous emotional appraisal.

This vision of the dual process of risk perception is part of the broader problem regarding the distinction between cognitive processes that are fast, automatic and unconscious (type 1) and those that are slow, laborious and conscious (type 2) (Evans, 2008; Evans and Stanovich, 2013). Risk as analysis would belong to type 2 processes, which rely on working memory and involve the mental simulation of future possibilities to formulate explicit judgements. In contrast, risk as feeling would belong to type 1 processes, which are autonomous, do not require working memory and underlie implicit information processing. However, as Evans (2008) rightly pointed out, the nature of the distinction between the two types of processes and their mutual relations are not univocal in the literature. This study addresses this issue through the prism of risk perception in a specific context, that of autonomous vehicles navigating shared spaces.

1.3. Risk Measurement

Herrero-Fernández et al. (2020) associated the concept of risk as analysis with a subjective assessment (SA) and associated the concept of risk as feelings with an objective evaluation based on an individual's physiological state. These two evaluation systems are complementary. If an SA consists of measuring the self-reflexive part of risk perception, then the physiological variables provide information on the physical manifestations of this same perception. From this perspective, Choi et al. (2019) considered that, at a certain level, risk perception may require the regulation of the autonomic nervous system, particularly by the sympathetic nervous system. The latter ensures that physiological adaptations occur in preparation for an escape or a struggle when a person is confronted with a stressful event. Such adaptations can manifest as increased cardiac and respiratory rhythms and variations in electrodermal activity (EDA). EDA corresponds to electrical variations in the skin that occur in relation to the functioning of the sweat glands, which are under the control of the sympathetic nervous system (Morange-Majoux, 2017). The most studied property of EDA is skin conductance, which is measured in micro-siemens and consists of the superimposition of two distinct parts that are called "tonic" and "phasic," respectively. The tonic component is associated with a global skin conductance level (SCL). It is relative to an individual and can be recorded when and individual is at rest. This component reveals slow variations, whereas the phasic component generally reveals rapid changes in skin conductance, which are often called "skin conductance responses (SCRs)." Choi et al. (2019) considered that risk perception could lead to substantial changes in EDA; therefore, EDA could be a good indicator of risk perception. SCRs have already been used as indicators of events that cause stress or discomfort in drivers. For example, Distefano et al. (2020) conducted an experiment using a driving simulator that revealed changes in the EDA of their participants as they approached intersections or roundabouts. Daviaux et al. (2020) observed similar effects when participants in their study were confronted with different driving situations (e.g., the insertion of another vehicle into the lane, crossing with another vehicle going in the opposite direction and crossing with a pedestrian). Skin conductance can be measured non-invasively using two electrodes that are placed on the surface of the skin (Fowles et al., 1981; Boucsein, 2012). In a detailed review about EDA, Boucsein (2012, p. 104–109) presented two preferable sites for placing the electrodes: The hand and the foot. However, the Society for Psychophysiological Research (Society for Psychophysiological Research *Ad Hoc* Committee on Electrodermal Measures, 2012) recommended placing the electrodes on the distal phalanges of the index and middle fingers to obtain bipolar recordings.

To investigate risk perception, this study proposes coupling this physiological measurement with a real-time subjective assessment by using an analogue device that can be operated using one hand. A similar method was used by Rossner and Bullinger (2019, 2020a,b). In their study, the participants were asked to assess their levels of comfort while they were on board a simulated autonomous vehicle. The advantage of collecting an online subjective measurement is that it provides access to the dynamics of changes in perceived risk, unlike verbal or written assessments that involve either an interruption in the task or an *a posteriori* evaluation. The device developed for this study resembles the slide potentiometer used by Walker et al. (2019a). In their study, the authors proposed an experiment to validate their measurement device by asking participants (pedestrians) to assess their willingness to cross a road in real-time while observing an approaching vehicle. The authors concluded that such a continuous assessment device could be useful for assessing human interactions with automated vehicles.

1.4. Theoretical Hypothesis

As mentioned previously, two factors that are related to vehicle-environment dynamics were manipulated: The value of the TTC at the moment the vehicle initiates an avoidance manoeuvre and the offset distance left between the vehicle and the pedestrian. It was postulated that the two factors would successively influence passengers' perceived risk. First, based on the results from Bootsma and Craig's study (Bootsma and Craig, 2003), it was assumed that perceived risk increases when the TTC at the time the manoeuvre is initiated decreases. Second, it was assumed that the closer the vehicle passes to a pedestrian (that is, the lower the offset distance), the greater the perceived risk becomes. Both TTC and offset were manipulated to investigate the relationship between the two types of risk perception measures (i.e., the subjective assessment and skin conductance response). From a statistical point of view, the objective was to determine whether the independence of the two types of risk perception was probable given the data and the effects of the independent factors. It was assumed that the measures of SA and SCR were continuous random variables. The purpose of this study was to test whether the experimental data would support the independence between those two variables given the levels of the factors being manipulated. Two alternatives were considered:

- $\mathcal{H}_0$: SA and SCR are independent despite the measures and levels of factors, which means that the two types of risk perception are independent.
- $\mathcal{H}_1$: SA and SCR are not independent in at least one combination of measures and factors, which means that there is a relationship between the two types of risk perception.

To address these theoretical hypotheses, hybrid Bayesian networks were implemented. This method, which is based on stochastic distributions, aided in discovering the best structure of relationships (i.e., the one that best fits the data) between manipulated factors and dependant measures. Specifically, Bayesian networks were used to determine whether a relationship between the two risk perception systems could exist and be relevant apart from the assumed influence of the independent environmental factors. In other words, this method was used to assess the significance of the relationship between measures from

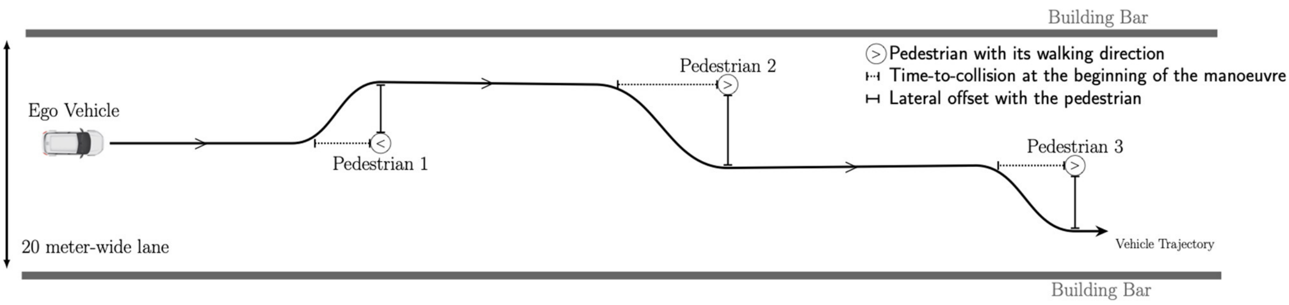


FIGURE 1 | Sample diagram of the vehicle trajectory among pedestrians. The combination of the two factors (TTC and offset) varied between pedestrians according to the experimental plan, which was identical for all the participants. The pedestrians' walking directions and the lateral directions of the avoidance manoeuvres are also represented.

two types of risk perception given the effects of the TTC and the lateral offset.

Finally, a network coefficient analysis was performed to quantify the effects of the manipulated factors. This analysis was conducted to test the extent to which the influence of the factors was confirmed. That is, this analysis was conducted to validate that a lower TTC at the beginning of an avoidance manoeuvre or a lower offset distance between a vehicle and a pedestrian results in a higher level of risk perception.

2. METHOD

2.1. Participants

For this experiment, 20 participants (13 men and 7 women) were recruited. They were between 18 and 52 years old ($M = 27.1$, $SD = 8.9$). Eighteen participants had held driving licenses for 9.8 years on average ($SD = 9.5$) and drove ∼11,400 km per year ($SD = 19{,}100$). The two remaining participants did not possess driving licenses, with the assumption that their perception of risk depended on the same processes as ordinary drivers. In addition, having a driver's license was not considered a prerequisite to be a passenger of a fully autonomous vehicle. Preliminary inspection of the EDA recordings and subjective assessments confirmed that the responses of the two unlicensed participants were not distinct from those of the others.

2.2. Experimental Design

A within-subject design was used so that participants would experience a series of 32 pedestrian avoidance maneuvers. The order of presentation of the pedestrians was randomized. They were divided into two blocks of ∼7 min presented in succession with a short break in between. The autonomous vehicle was programmed to drive at a constant speed of 30 kph on a 20-m-wide street. The vehicle followed a straight trajectory except when it had to avoid pedestrians. In the real world, the speed limit in shared space is generally lower. According to the British Department of Transport (Great Britain and Department of Transport, 2011), a speed of no more than 20 mph (∼32 kph) and preferably <15 mph (∼24 kph) is desirable. However, some preliminary experiments revealed that at low vehicle speeds, avoidance manoeuvres elicit very little risk perception from participants. This phenomenon may be explained by the relatively limited immersion of the driving simulator and participants' ability to predict the behavior of the autonomous vehicle. Therefore, the speed of 30 kph was chosen to increase the chances of eliciting risk perception from the participants.

A sample of the vehicle's trajectory is illustrated in **Figure 1**. After passing a pedestrian, the vehicle did not return to its initial position but maintained its position in the lane until the next pedestrian was encountered. Each pedestrian walked at 4.5 kph and was 25 s apart from the others.

2.3. Factors

Two factors acting successively were manipulated (cf. **Figure 1**).

First, the value of the TTC when the avoidance manoeuvre was initiated was manipulated. In a straight line, the TTC represented the time remaining before the vehicle reached an obstacle. This depended on both the distance and the relative speed between the vehicle and the obstacle. During the experiment, four levels of TTC were tested: 2.0, 2.5, 3.0, and 3.5 s.

Second, the lateral offset distance (simply denoted "the offset") was manipulated during ongoing avoidance manoeuvres when the vehicle arrived next to a pedestrian. This parameter was introduced to test whether the proximity between the vehicle and the pedestrian affected the participant's perceived risk. Three levels of lateral offset were tested: 0.5, 1.0, and 1.5 m.

It was not possible to combine all the levels of the two factors. Indeed, combining a time-to-collision of 2.0 s and an offset of 1.0 or 1.5 m gave rise to unrealistic vehicle behavior that was caused by the driving simulation software. As a result, only 8 out of the 12 combinations were tested.

Two additional factors were introduced to make the simulations more realistic and unpredictable. Half of the pedestrians walked in the direction in which the vehicle was moving, while the other half walked in the direction opposite the vehicle. In a shared space, there are no rules regarding the direction in which a vehicle should go to avoid other road users. For this reason, the direction of the avoidance manoeuvres varied between left and right. Preliminary statistical analyses, which have not been reported here, demonstrated that these two factors did not affect the results. Finally, the appearance of

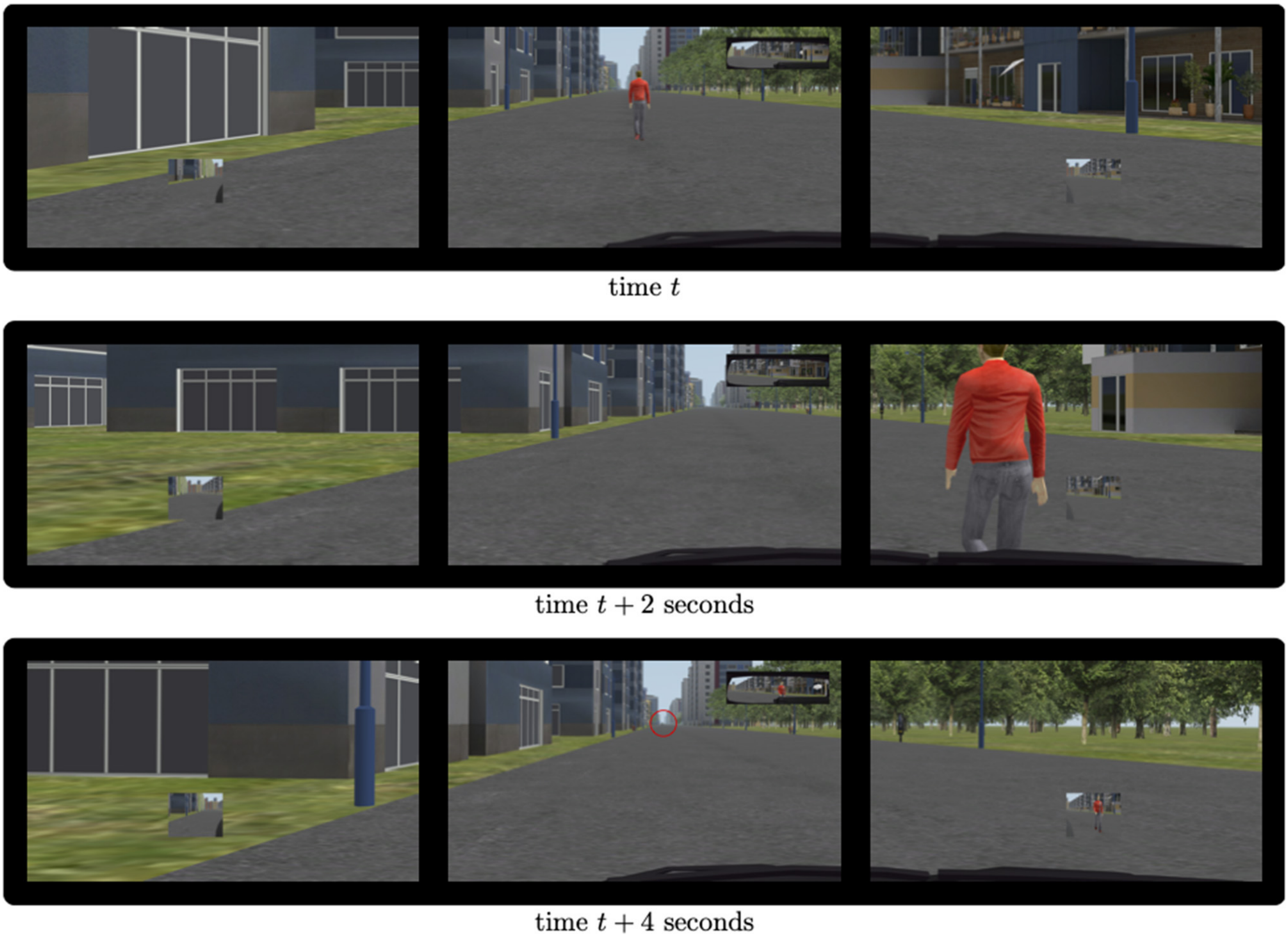


FIGURE 2 | Example of the participant's 120° view during a left avoidance manoeuvre. **(Top)** The approaching phase of a pedestrian walking in the same direction as the vehicle. **(Middle)** The vehicle's position 2 s later when it had started its avoidance manoeuvre. **(Bottom)** The end of the manoeuvre. The avoided pedestrian is visible in the left and center rear-view mirrors. The following pedestrian (circled in red) can hardly be distinguished. This is because each avoidance manoeuvre was separated by 25 s to ensure the independence of the measurements.

each pedestrian was arbitrarily chosen from a list of a dozen possibilities (a man or a woman in a t-shirt or a suit, a teenager in shorts, etc.). **Figure 2** illustrates what the participants saw during a left avoidance manoeuvre in three screenshots.

2.4. Experimental Setup

The experiment took place using a fixed-base driving simulator that was run using SCANeR Studio™ software (AVSimulation, France). This driving simulator provides visibility of 120° thanks to three screens (see **Figure 3A**). During the simulation, the participants were informed that the vehicle was fully automated and that they did not need to use the controls.

In order to record physiological responses during the simulation, the participants' skin conductance was measured according to recommendations of the Society for Psychophysiological Research (Society for Psychophysiological Research *Ad Hoc* Committee on Electrodermal Measures, 2012, p. 6-7). Two electrodes were placed on the distal phalanges of the participants' index and middle fingers on their non-dominant hands. To improve the skin-electrode electrical conductivity and the accuracy of the data collection, the electrodes were covered with isotonic gel. No skin preparation was done before the electrodes were placed. Data recording began at least 5 min after establishing electrode-skin contact to create better electrical contact and stabilize the baseline (that is, the skin conductance level). The data were collected at 625 Hz on a dedicated computer using the software AcqKnowledge 5.0 (BIOPAC Systems, Inc., USA), which was coupled with an acquisition module (16 bits analogue to digital converter; MP160 system, Systems, Inc., USA).

To enable the participants to assess perceived risk throughout the simulation, an analogue device (a potentiometer that was connected to an Arduino*™* Uno board) was developed for one-handed use, which is illustrated in **Figure 3B**. The device was designed so that it did not cause visual distraction and could be used without the participants having to look at it. The device was placed on the participants' laps in such a way that they could manipulate it using their dominant hands (i.e., the hand that did

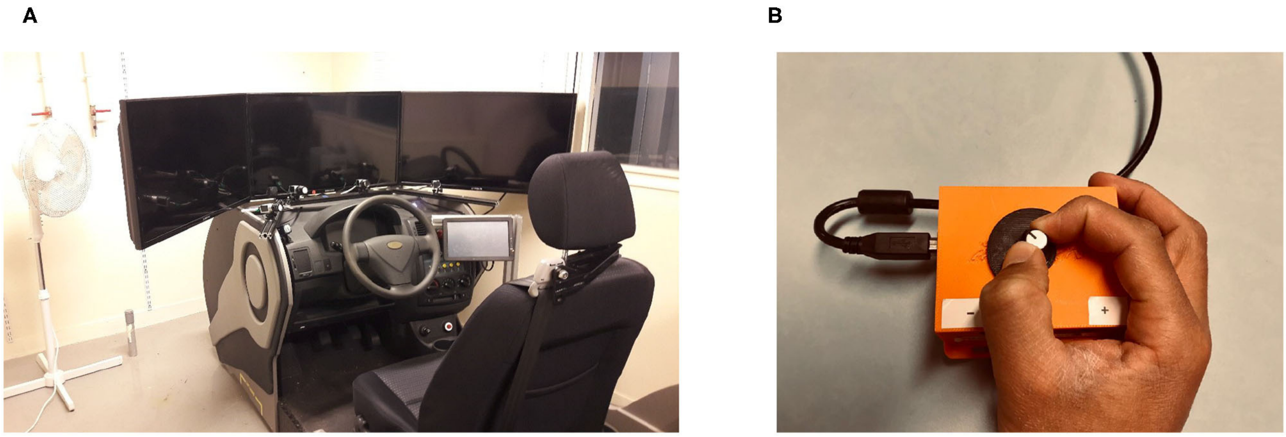


FIGURE 3 | Experimental setup. **(A)** A fixed-base driving simulator. **(B)** Analogue device for subjective risk assessment.

not have the conductance measurement electrodes attached to it). Data were collected and linked to the autonomous driving scenarios with a sampling rate of 20 Hz.

2.5. Procedure

Participants were asked to fill out a form to provide information about their ages, genders, and driving experience. They were then informed of the purpose and the course of the experiment. At the end of the introductory period, the electrodes were installed according to the manual preferences of the participants. Each participant was then invited to sit in the simulator; at this point, the analogue device was presented and handed over to them.

Each participant was presented with a preliminary scenario that consisted of autonomously driving on a road without pedestrians. Each participant's objective was to optimize the use of the hand-held device by finding a good position for their hands and exploring the rating scale available. Voluntarily, no scale or reference value that was related to risk assessment was provided to the participants. Therefore, they had no prior knowledge of the lowest or highest levels of stimulation that they would encounter. This approach was inspired by Stevens' book about psychophysics (Stevens, 2017, p. 28): This testing method gave participants more freedom without distracting them as it did not require them to do calculations to scale their responses to a certain criterion. The participants were expected to pay attention as much as possible to the driving scene and ignore the assessment device. During this scenario, a horizontal gauge indicating the cursor's position in real-time was displayed on the central monitor. The gauge represented the position of the cursor in the usual way (that is, with the minimum value on the left and the maximum value on the right). The participants were initially invited to adjust their positions in the driver's seat and to familiarize themselves with the assessment device without the researcher's intervention. They were then asked to perform a few exercises: Starting from either the cursor's minimum or maximum position, they had to reach the first third, the median and then two-thirds of the gauge. During this training phase, the participants were asked to close their eyes and reopen them when they thought they had reached the required positions. In this way, they could estimate their errors and readjust their positions if necessary. An error of 5% around the target position was allowed. This phase ended as soon as the participants managed to reach all the required positions and felt confident enough to reach any other position. This training period also made it possible to check the quality of the physiological data collected (EDA) and to adjust the electrodes' placement if necessary.

Afterwards, the participants were instructed to experiment "using the analogue device to assess their risk of collision with pedestrians in real-time when moving in a shared space."

2.6. Calculation of Dependent Variables

2.6.1. Subjective Risk

Two indicators were calculated to quantify the subjective risk assessment (denoted as "SA") for each avoidance manoeuvre: The integrated risk assessed over time (*iSA*) and the maximum amplitude of the assessed risk (*mSA*). The indicator *iSA* can be considered a dynamic indicator as it accounts for both amplitude and temporality. It corresponds to the area under the curve for each maneuver filled with grey in **Figure 4**. The indicator *mSA* was calculated to provide a simple quantification of the participants' SA. It corresponds to the maximum value at each peak. As illustrated in **Figure 4**, the subjective assessment evolved differently between the manoeuvres and always returned to 0 between each pedestrian.

2.6.2. Skin Conductance Response

The data were initially processed using the software program AcqKnowledge 5.0 (BIOPAC Systems, Inc., USA); MATLAB (MATLAB, 2018) and R (R Core Team, 2020) were then used to extract the indicators. Several manipulations had to be conducted to calculate the indicators. First, the raw data were pre-processed using AcqKnowledge according to the recommendations made by Braithwaite et al. (2013) and Findlay (2017). These included resampling at 78 Hz, moving median smoothing at 1 s and low-pass filtering at 1 Hz. The pre-processed data were then analyzed using the Ledalab application, specifically Version 3.4.8

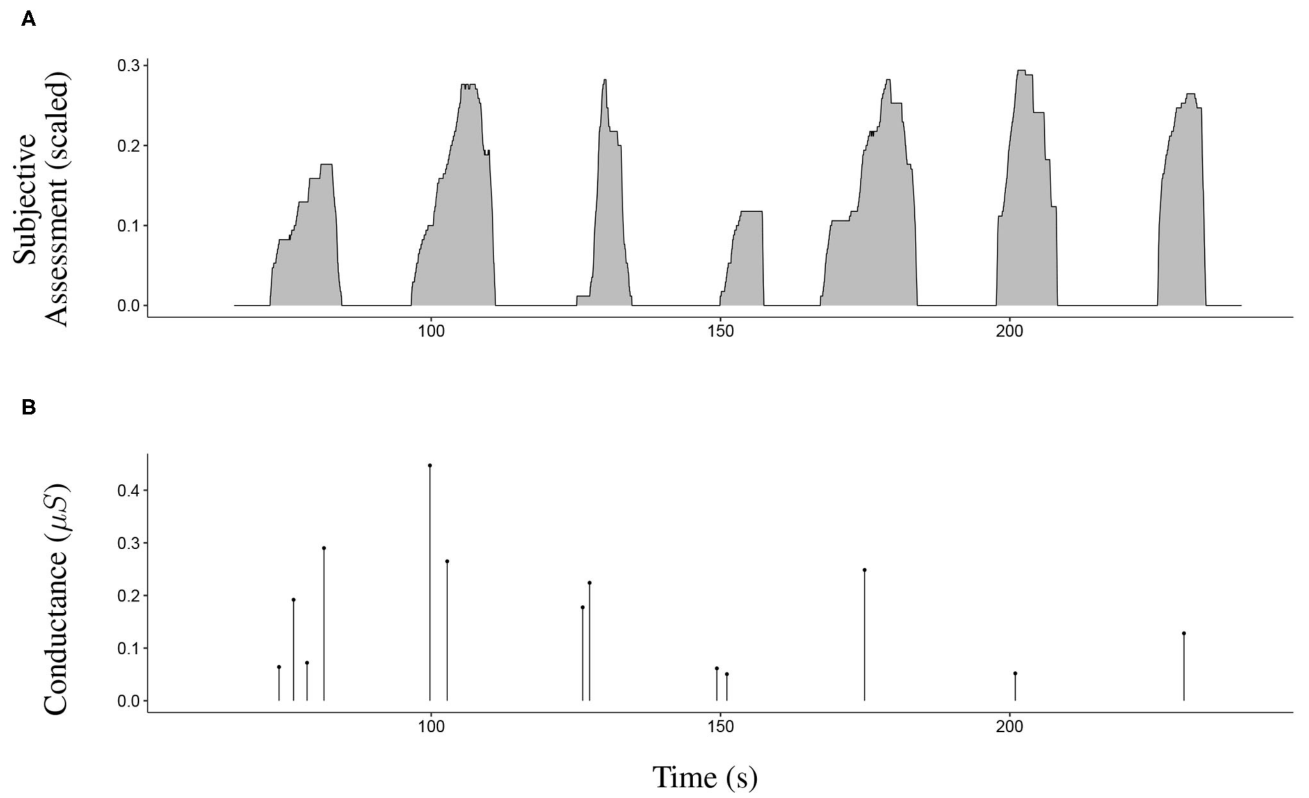


FIGURE 4 | Sample of the collected data concerning seven avoidance manoeuvres. **(A)** The evolution of the subjective assessment during seven avoidance manoeuvres. The subjective assessment always returned to 0 between the manoeuvres. Indicator mSA corresponds to the maximum value of each manoeuvre. Indicator iSA corresponds to the area under each curve. **(B)** The skin conductance responses during the seven avoidance manoeuvres. Indicator mSCR corresponds to the maximum amplitude of the responses. Indicator nSCR corresponds to the number of skin conductance responses.

from Benedek and Kaernbach (2010). This application is a MATLAB toolbox that was designed for isolating the tonic and phasic components of EDA. The method used was based on the following two steps:

1. The identification of the phasic component using Continuous Decomposition Analysis (CDA) with two parameters for the optimization of the deconvolution algorithm. This method, which is described in Benedek and Kaernbach's article (Benedek and Kaernbach, 2010), is based on a deconvolution algorithm to detect SCR. This technique is particularly effective for identifying and determining the characteristics of so-called "superimposed" responses (Boucsein, 2012).
2. The detection of SCR > 0.05 μS (the extraction of their onset and amplitude).

Finally, the indicators were calculated using the R software. For each avoidance manoeuvre, relative SCRs were selected only if their onset occurred no more than w_{start} seconds before the moment when the vehicle was next to a pedestrian and no more than w_{end} seconds after that moment. An avoidance manoeuvre could elicit an SCR only during the moment the participants started to perceive (assess) a collision risk. Therefore, the value w_{start} was computed for each participant, and two distinct moments for each manoeuvre were considered:

- The moment when the participant started to assess a non-zero value of collision risk;
- The moment when the vehicle was effectively level with a pedestrian.

This process resulted in 32 values (which corresponded to the total number of manoeuvres) per participant, which were averaged to find w_{start}. The value of w_{end} was based on results from the literature (Boucsein, 2012; Droulers et al., 2013). SCR could be related to an avoidance manoeuvre only if it occurred in the 3 s following the moment when the vehicle was next to a pedestrian.

As for the subjective assessment data, two indicators were calculated to quantify EDA. These included the maximum amplitude of skin conductance responses (*mSCR*) and the number of skin conductance responses (*nSCR*). **Figure 4** illustrates a sample of SCR data. The indicator *nSCR* corresponds to the number of SCRs during each manoeuvre, and *mSCR* corresponds to the maximum amplitude for all concerned SCRs.

TABLE 1 | Value's frequencies for independent variables.

Offset (m)	TTC (s)				Total frequency
	3.5	3	2.5	2	
1.5	0.125	0.125	0.125	0.125	0.500
1	0.125	0.125	0	0	0.250
0.5	0.125	0.125	0	0	0.250
Total frequency	0.375	0.375	0.125	0.125	1.000

This table illustrates the combinations of independent variables used in the experiment. The frequency for each value is given, and the totals for each row/column were added to provide details on the distribution of each variable.

TABLE 2 | Sample descriptive statistics and normality test of subjective assessment variables.

Variable	Descriptive statistics				Shapiro Wilk test	
	n	M	SD	Skewness	W	*p*
Raw						
mSA	640	0.401	0.252	0.653	0.951	0.000
iSA	640	1.548	1.426	1.431	0.865	0.000
Transformed						
$(\text{mSA})^{\frac{1}{2}}$	640	0.000	0.985	−0.010	0.999	0.974
$(\text{iSA})^{\frac{1}{3}}$	640	0.000	0.985	−0.016	0.998	0.701

The raw variables correspond to the original data. The transformed data correspond to centered and scaled variables for each individual. These operations were performed after the power transformations of the initial values were made (that is $\frac{1}{2}$ for the indicator mSA *and $\frac{1}{3}$ for the indicator* iSA*).*

3. DATA ANALYSIS

The analysis was based on the modeling of hybrid Bayesian (Denis and Scutari, 2015, Chapter 3) to study the relationship that may exist between subjective risk assessment indicators and the participants' skin conductance response indicators and the effects of the factors. Each Bayesian network proposed in this study includes four variables, which are also called "nodes." A Bayesian network is represented by its directed acyclic graph (DAG), which graphically illustrates the relationships between its nodes. A node represents a variable that is associated with a statistical distribution whose parameters possibly depend on the other nodes. An arrow is used to specify that the distribution of a node depends on the value of another node. In this study, all directed acyclic graphs contain four nodes. Two nodes represent the independent factors TTC and offset. They were both assigned to discrete distributions whose probability mass functions were determined by the frequency of their values in the design of the experiment (cf. **Table 1**). Two other nodes, which were denoted "SA" and "SCR," were respectively assigned to the subjective assessment and the skin conductance response indicators. Node SA (resp. SCR) designated either the *iSA* indicator or the *mSA* indicator (resp. *nSCR* and *mSCR* indicators).

After an analysis of the distributions of the two indicators of subjective assessment was conducted, the original data was transformed to correct a positive skewness. A power of $\frac{1}{2}$ was applied to *mSA* values and a power of $\frac{1}{3}$ to *iSA* values. Moreover, to consider global distributions for all the participants, the transformed SA values were then centered and scaled by the participants. These transformations were performed to use Gaussian distributions for the node SA in the Bayesian networks. To ensure that this hypothesis on the distributions was relevant, a Shapiro-Wilk test was performed. The results are presented in **Table 2**. The transformations that were performed on the indicators resulted in more symmetrical distributions that can be assumed to be normal according to the statistics of the Shapiro-Wilk test ($p > 0.5$ for both variables).

A preliminary analysis revealed that the SCR indicators had 46% of exactly zero. That means that only a part of an avoidance manoeuvre produced physiological responses. For this reason, a Tweedie compound Poisson distribution (Dunn and Smyth, 2005, 2008; Hasan and Dunn, 2012) was used for Node SCR in the Bayesian networks. This otherwise positive and continuous distribution has a positive mass at zero. The Tweedie compound Poisson distribution aided in estimating the distribution of the SCR indicators, as well as the probability of zero responses. To consider global distributions and homogenize the fluctuations between the data of each participant, indicators mSCR and nSCR were scaled per participant. As in the example provided by Dunn and Smyth (2005), an initial diagnostic (which has not been reported here) was performed to verify that the Tweedie approach of modeling the zeros and the positive observations together was adequate to estimate the parameters of the distribution.

For all node distributions, the parameters were estimated using the R software. More specifically, as in the method presented by Denis and Scutari (2015), the parameters of the factors TTC and offset were set as the actual frequency in the experiment (cf. **Table 1**), and the parameters of the Gaussian distribution were estimated by fitting linear models. The parameters of the compound Poisson distribution were estimated using the R package *cplm* (Zhang, 2013). Following the method presented by Denis and Scutari (2015), when a factor influenced a dependent node's (SA or SCR) distribution, the parameters were estimated for each value of the factor. For instance, eight parameters were estimated for the distribution of a Gaussian node that was influenced exclusively by the factor TTC (i.e., a mean and a standard deviation for each of the four levels of TTC). Concerning the graphs where both the factor TTC and offset influenced an indicator, a distribution was fitted for each of the eight combinations (cf. **Table 1**).

Forty-eight networks were computed regarding the four indicators mentioned previously (*iSA*, *mSA*, *nSCR*, and *mSCR*; see **Figure 5**). The consideration of four indicators rather than two (i.e., one for the SA and one for the SCR) allowed the amount of data that was used to analyze the effect of the factors and the relationship between the two risk perception types to be multiplied by four. To compare the Bayesian networks and select the more plausible one given the data, the Bayesian Information Criterion (BIC, Schwarz, 1978; Kass and Raftery, 1995; Raftery, 1995) was used. This is the criterion that is used for Bayesian network selection in the greedy search algorithm mentioned by Denis and Scutari (2015, p. 110). The procedure consists of favoring the network that has the lowest BIC score. For a Bayesian network, the Bayesian Information Criterion was calculated according to the following formula:

$$\mathbf{BIC} = -2\mathcal{LL} + p\log(n), \tag{1}$$

where $\mathcal{LL}$ is the joint log-likelihood of variables in the Bayesian network, p is the number of estimated parameters associated with the joint distribution of the variables in the network and n is the number of samples. The BIC allows non-nested Bayesian networks to be easily compared and is conservative regarding relationships (Raftery, 1999). That is, new relationships between nodes will only be significant concerning the BIC if they provide sufficient benefits regarding the overall likelihood. In the specific case of the Bayesian network, the BIC calculation was decomposed as the sum of the BIC at the four nodes (Denis and Scutari, 2015, p. 19):

$$\mathbf{BIC} = \mathbf{BIC}^{\mathrm{TTC}} + \mathbf{BIC}^{\mathrm{Offset}} + \mathbf{BIC}^{\mathrm{SA}} + \mathbf{BIC}^{\mathrm{SCR}}, \tag{2}$$

where $\mathbf{BIC}^x$ is the value of the Bayesian information criterion computed for a node x. This equation results from the fact that the joint log-likelihood of a Bayesian network can itself be decomposed as the sum of the log-likelihood for each node when considering the relationships between them while estimating distribution parameters.

The grades of evidence from Kass and Raftery (Raftery, 1995, p. 139) were used to discuss the BIC differences in the values after the ranking process. Gaps larger than 2, 6, or 10 between two BIC values were considered positive, strong or very strong, respectively. Afterwards, the Hypothesis $\mathcal{H}_0$, which states the independence between the subjective assessment and the skin conductance responses, requires that the best Bayesian networks (which were obtained for each combination of indicators) do not contain a relationship between the node SA and the node SCR. Conversely, it is sufficient for one of the best networks to contain a relationship between the node SA and the node SCR to reject Hypothesis $\mathcal{H}_0$ in favor of $\mathcal{H}_1$.

The best Bayesian networks were finally investigated to analyse the estimated distribution of each indicator. For indicators of subjective assessment, whose distributions were assumed to be Gaussian, the conditional mean estimates with confidence interval at 95% were represented. Additionally, when the influence of both factors appeared in the best Bayesian network, a cluster analysis was performed based on the Bayesian information criterion (Binder, 1978; Franzén, 2008). We considered all the conditions resulting from the interaction of the TTC and offset factors, that is, eight possible levels. The objective was then to find out if these eight levels gave rise to different distributions of the indicator considered (*mSA* or *iSA*) or if they could be grouped into a smaller number. For this purpose, models were built for all possible groupings, that is, 4140 possible partitions in accordance with the eighth Bell number (Rota, 1964). The R package *partitions* was used for this (Hankin, 2006; Hankin and West, 2007). Then the BIC value was calculated for all the models. The best model was

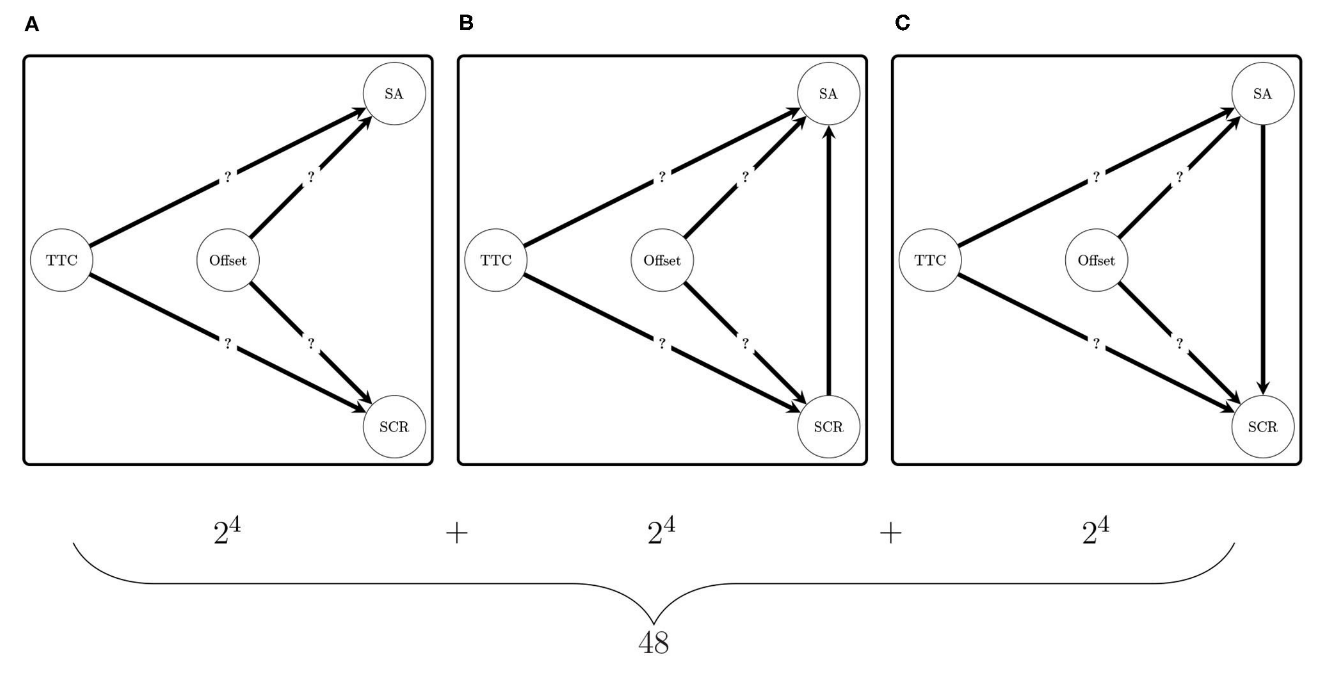


FIGURE 5 | Directed acyclic graphs (DAG) compared in the experiment. **(A)** Illustrates the 16 DAGs that do not contain a relationship between SA and SCR nodes. **(B)** Illustrates the DAG that contains a relationship between the node SCR and the node SA. **(C)** Illustrates the DAG that contains a relationship between the node SA and the node SCR. The only DAG in **(A)** is consistent with the hypothesis of independence between the subjective assessment and the skin conductance response.

selected by using the Raftery's grades of evidence (Raftery, 1995, p. 139).

For indicators of SCR, which were assumed to follow Tweedie distributions, conditional mean estimates were represented, as well as the probability of zeros. These two representations were used to provide a more complete preview of the functioning of an effect given this specific Tweedie distribution. Both representations were useful for providing a more complete picture of how an effect worked given this Tweedie distribution. Since SCRs were quite rare in the data, it was interesting to visualize the evolution of the mean of an indicator in parallel with the probability of no response.

4. RESULTS

4.1. The Relationship Between the Two Types of Risk Perception Measures

All the possible Bayesian networks were compared to address the two theoretical hypotheses: $\mathcal{H}_0$, which states that the two types of risk perception are independent, and $\mathcal{H}_1$, which states that a relationship exists between the two types of risk perception. The best Bayesian networks were selected for each of the four combinations of indicators according to the BIC. **Table 3** presents the three best Bayesian networks per combination regarding their value of BIC. In this table, the structure of the relationships between the Bayesian networks was represented by the likelihood decomposition for node SA and node SCR. This notation was adopted to succinctly reflect the dependence between the distributions of subjective assessment measures and skin conductance responses given the factors TTC and offset.

TABLE 3 | Values of the BIC of the three best models by indicator combination.

Likelihood decomposition		BIC
$\mathcal{L}_{(\text{mSA, mSCR}\mid\text{TTC, Offset})}$		
$\mathcal{L}_{(\text{mSA}\mid\text{TTC})}$	$\times\ \mathcal{L}_{(\text{mSCR}\mid\text{TTC})}$	6044.724*
$\mathcal{L}_{(\text{mSA}\mid\text{TTC, Offset})}$	$\times\ \mathcal{L}_{(\text{mSCR}\mid\text{TTC})}$	6045.361*
$\mathcal{L}_{(\text{mSA}\mid\text{TTC, mSCR})}$	$\times\ \mathcal{L}_{(\text{mSCR}\mid\text{TTC})}$	6050.928
$\mathcal{L}_{(\text{mSA, nSCR}\mid\text{TTC, Offset})}$		
$\mathcal{L}_{(\text{mSA}\mid\text{TTC})}$	$\times\ \mathcal{L}_{(\text{nSCR}\mid\text{mSA})}$	5979.376*
$\mathcal{L}_{(\text{mSA}\mid\text{TTC, Offset})}$	$\times\ \mathcal{L}_{(\text{nSCR}\mid\text{mSA})}$	5980.013*
$\mathcal{L}_{(\text{mSA}\mid\text{TTC})}$	$\times\ \mathcal{L}_{(\text{nSCR}\mid\text{TTC})}$	6009.594
$\mathcal{L}_{(\text{iSA, mSCR}\mid\text{TTC, Offset})}$		
$\mathcal{L}_{(\text{iSA}\mid\text{TTC, Offset})}$	$\times\ \mathcal{L}_{(\text{mSCR}\mid\text{TTC})}$	6038.815*
$\mathcal{L}_{(\text{iSA}\mid\text{TTC})}$	$\times\ \mathcal{L}_{(\text{mSCR}\mid\text{TTC})}$	6053.426
$\mathcal{L}_{(\text{iSA}\mid\text{TTC, Offset})}$	$\times\ \mathcal{L}_{(\text{mSCR}\mid\text{TTC,iSA})}$	6057.045
$\mathcal{L}_{(\text{iSA, nSCR}\mid\text{TTC, Offset})}$		
$\mathcal{L}_{(\text{iSA}\mid\text{TTC, Offset})}$	$\times\ \mathcal{L}_{(\text{nSCR}\mid\text{iSA})}$	5983.842*
$\mathcal{L}_{(\text{iSA}\mid\text{TTC})}$	$\times\ \mathcal{L}_{(\text{nSCR}\mid\text{iSA})}$	5998.453
$\mathcal{L}_{(\text{iSA}\mid\text{TTC, Offset})}$	$\times\ \mathcal{L}_{(\text{nSCR}\mid\text{TTC})}$	6003.685

For the sake of simplicity, the likelihood (denoted $\mathcal{L}$) decompositions are only noted for nodes SA and SCR. However, the values in the BIC column correspond with the total BIC of the Bayesian network. The distribution parameters have not been specified here. The indicators of subjective assessment were assumed to follow a Gaussian distribution. The indicators of skin conductance response were assumed to follow a Tweedie distribution. The asterisks indicate the best Bayesian networks according to Raftery's degree of evidence.

Table 3 reveals that one unique Bayesian network was selected for the indicator *iSA*. The results revealed that this indicator was influenced by the combination between levels of TTC and offset. Concerning the indicator *mSA*, the best Bayesian network (i.e., the Bayesian network with the lowest BIC score) could not be definitely distinguished from the Bayesian networks ranked in second position. The BIC difference between those two best Bayesian networks was not significant regarding the Raftery's grade of evidence (Raftery, 1995). Essentially, the difference was lower than 2. The dependence structure of those two best Bayesian networks were similar except for the dependence between the indicator *mSA* and the factor offset, which appeared only in the Bayesian network that was ranked in second position. This result means that there is not enough evidence to conclude with certainty that the factor offset influenced the indicator *mSA*.

The directed acyclic graphs of the best Bayesian networks revealed by **Table 3** are illustrated in **Figure 6**. Details about the estimated coefficients for each distribution are provided in **Figure A1**. Since there was not enough evidence to conclude the relationship between the factor offset and the indicator *mSA*, a question mark was placed on the arrows between these two nodes. **Figure 6** illustrates two results.

- First, the subjective risk assessment depends on the two factors: Indicator *mSA* definitely depends on the factor TTC and possibly on the factor offset (see **Figures 6A,C**), and indicator *iSA* depends on both factors TTC and offset (see **Figures 6B,D**).
- Secondly, whereas the maximum amplitude of the skin conductance responses (indicator *mSCR*) depends only on the TTC (see **Figures 6A,B**), the number of skin conductance responses depends only on the subjective risk assessment (see **Figures 6C,D**). This result permitted the rejection of $\mathcal{H}_0$ in favor of $\mathcal{H}_1$.

Consequently, the results support the existence of a relationship between the two types of risk perception.

4.2. The Analysis of Risk Perception Variations According to the Factors

A coefficient analysis was performed for the subjective risk assessment obtained for each observed combination of the factors. **Figure 7** presents the means of the subjective risk assessment obtained for the indicators *mSA* (**Figure 7A**) and *iSA* (**Figure 7B**). The means were typically represented by bars that were proportional to their value, and 95% confidence intervals were plotted to better visualize significant differences. An inspection of this figure reveals that the patterns of the results (as a function of the combinations of factor) were similar for both indicators. A cluster analysis was performed to better identify the similarities and differences between the means of a given indicator. This procedure relied on the Bayesian information criterion to find the optimal groupings of factor level combinations based on the data. It resulted in three

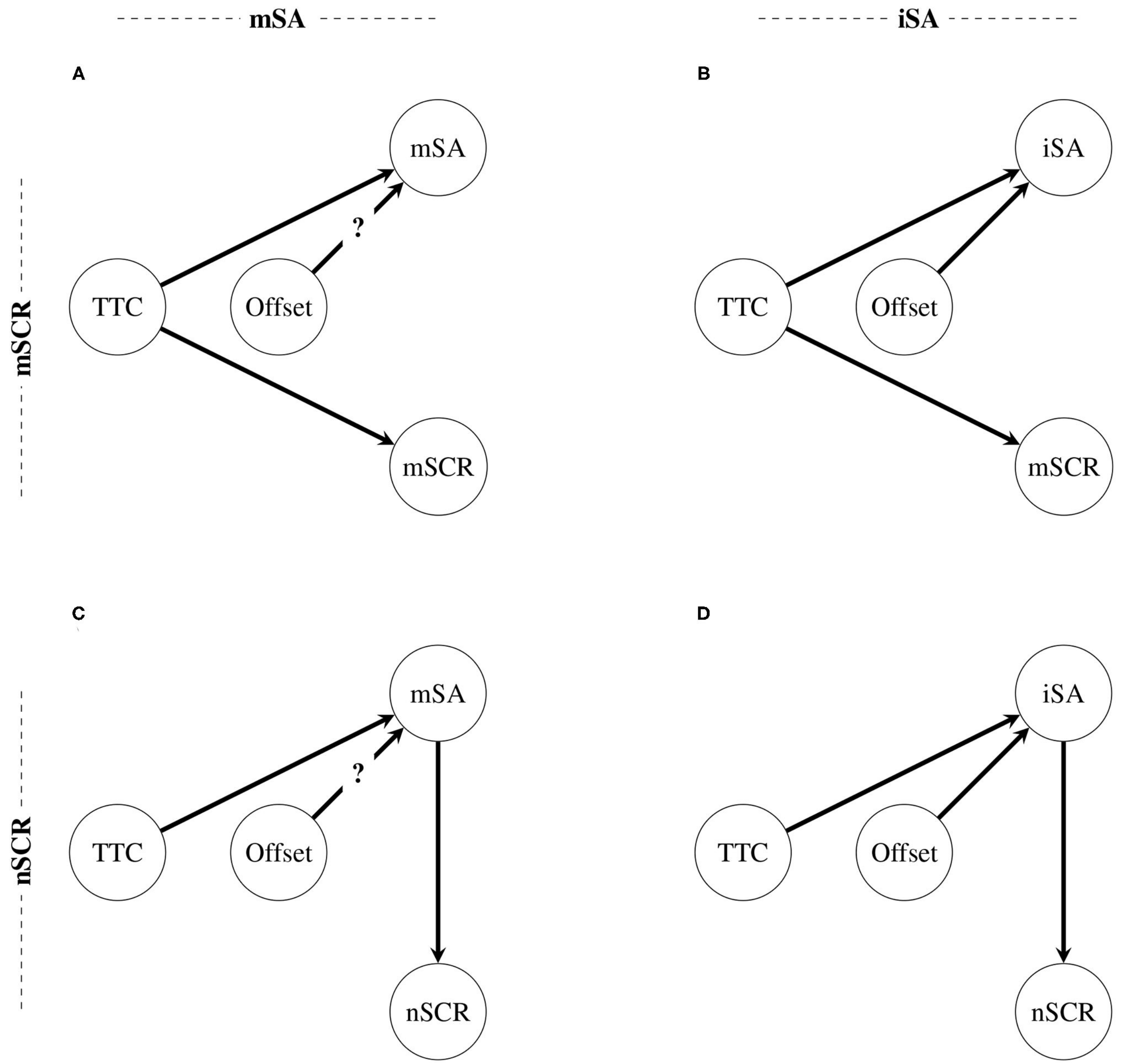


FIGURE 6 | Directed acyclic graphs of the best Bayesian networks by indicators combination. **(A)** The relationships between the factors and the indicators *mSA* and *mSCR*. **(B)** The relationships between the factors and the indicators *iSA* and *mSCR*. **(C)** The relationships between the factors and the indicators *mSA* and *nSCR*. **(D)** The relationships between the factors and the indicators *iSA* and *nSCR*. A question mark was introduced to the illustration of relationship between node Offset and node *mSA* because the degree of evidence for this relationship is not sufficient regarding the BIC given the thresholds developed by Raftery (1995), which were adopted in this study.

homogeneous categories that were characterized as functions of the effect on risk perception: Low, mid, and high. These three categories have been detailed in the following manner:

- When the values of the factors TTC and offset were both high (superior or equal to 3.0 s for the TTC and superior or equal to 1.0 m for the offset), the risk was perceived to be low; the means of the subjective assessment indicators were lower than the average.
- When the level of the TTC was median (2.5 s) or when the offset was small (0.5 m), the risk was perceived to be moderate; the means of the subjective assessment indicators were close to the average.
- When the level of the TTC was small (2.0 s), the risk was perceived to be high; the means of the subjective assessment indicators were significantly greater than the average.

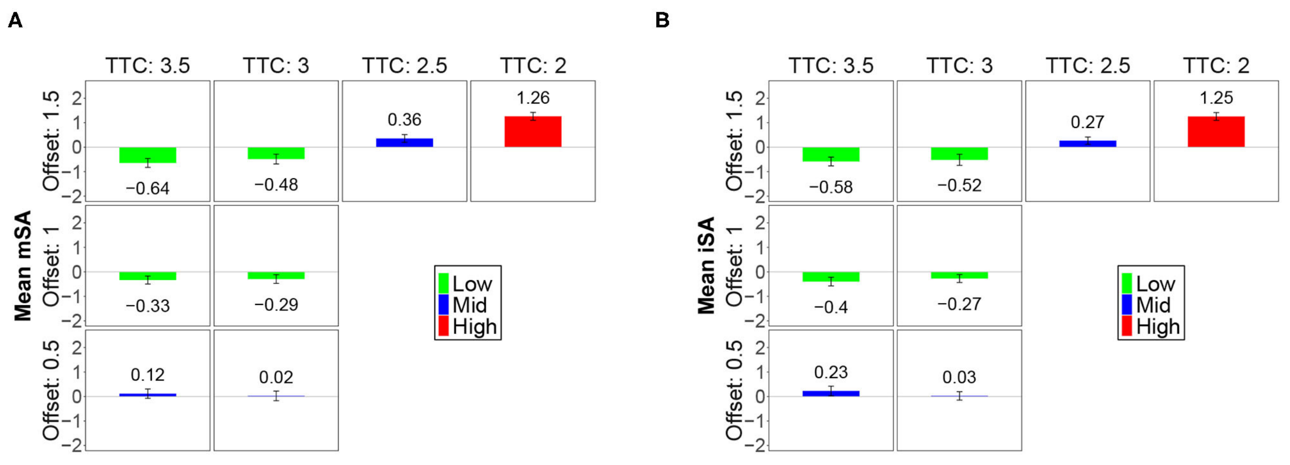


FIGURE 7 | Means for the SA variables for level combinations of TTC and offset. **(A)** The means of the indicator *mSA* in combination with the TTC and offset. **(B)** The means of the indicator *iSA* in combination with the TTC and offset. To better distinguish differences, confidence intervals at 95% are represented for each mean.

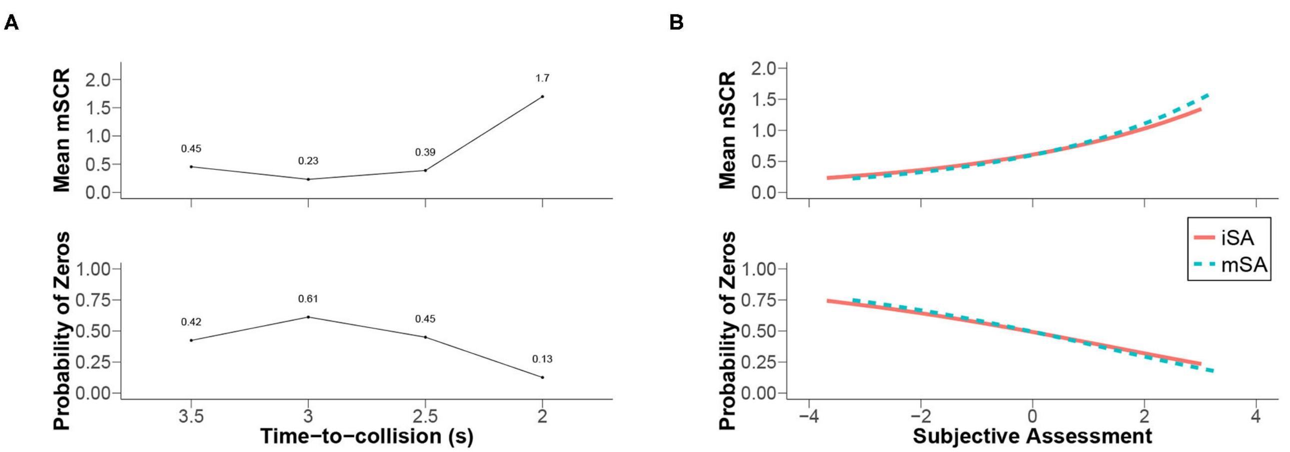


FIGURE 8 | Conditional means of SCR indicators according to the best Bayesian networks. **(A)** The means of the indicator *mSCR* in combination with the TTC values. **(B)** The means of the indicator *nSCR* as a function of the SA values. The probability of zeros was derived from the estimated parameters of each distribution (Tweedie). The relationships between the parameters of such distributions were, for instance, presented by Zhang (2013, eq. 2.2).

Figure 8A presents the estimated means of the indicator *mSCR* and the estimated probabilities of zeros according to the levels of factor TTC. The examination of this figure revealed that the probability of zeros (i.e., of not observing skin conductance responses in the participants) decreased as the mean of the indicator *mSCR* increased. Moreover, the maximum amplitude was obtained when the level of the TTC was small (2.0 s). Consequently, the later the vehicle initiated its avoidance manoeuvre, the greater the chance of observing skin conductance responses in the participants became. Furthermore, it can be noted that the lowest level of the TTC produced a similar impact on the two risk perception systems. In this particular case, the risk was perceived to be high based on the indicators of the two types of risk perception.

The analysis of the directed acyclic graphs of the best Bayesian networks (cf. **Figure 6**) revealed that the number of skin conductance responses depended on the subjective risk assessment indicators (i.e., *mSA* and *iSA*) rather than on the TTC and offset levels. This result was considered an example of the relationship between the two types of risk perception as it refuted their independence. **Figure 8B** illustrates this relationship (using estimated parameters of the Tweedie distribution assumed for the indicator *nSCR* detailed in **Figure A1**). The mean of the number of skin conductance responses was represented as a function of the subjective risk assessment indicators. The mean of the number of skin conductance responses varied similarly for indicators *mSA* and *iSA*. Per the statistical model used to fit the distribution, the link between the number of skin conductance responses and the indicators of subjective risk assessment was exponential. The higher the subjective assessment indicators were, the higher the number of skin conductance responses were. Additionally, the probability of having no skin conductance response (i.e., the probability of zeros) was high when the indicators of subjective risk assessment were low. The probability of not observing a skin conductance response in the participants decreased as the subjective assessment indicators decreased.

5. DISCUSSION

This study sought to characterize the perception of risk made by a passenger in an autonomous vehicle that was moving in a space shared with pedestrians. For this purpose, the subjective risk assessment and the skin conductance responses were collected in parallel to better understand how the two perception systems ("risk as feeling" and "risk as analysis") act in such a situation.

The result of the Bayesian network modeling revealed that the hypothesis concerning independence between the two risk perception systems must be rejected under the TTC and offset conditions that the study evaluated. Although the maximum amplitude of the skin conductance responses is impacted by small TTC values, the analysis demonstrated that the number of skin conductance responses depends only on the subjective risk assessment. These results, therefore, support the hypothesis that claims that the two risk perception systems are not completely interdependent as they may influence one another independently of environmental factors (Loewenstein et al., 2001; Slovic et al., 2004). Nevertheless, since the subjective risk assessment was more sensitive to external conditions than the skin conductance responses, it is more likely that the subjective risk can induce skin conductance responses than the opposite. This conclusion will have to be confirmed by further studies.

The results revealed that there are three classes of situations. When TTC and offset were simultaneously high, between 3.0 and 3.5 s and between 1.0 and 1.5 m respectively, the risk was perceived as low. When the TTC was intermediate (2.5 s) or when the offset was low (0.5 m), the risk was perceived as moderate. Finally, when the TTC was small (2.0 s), the perceived risk was higher than in all other situations. Thus, the results confirmed that the TTC strongly determines the perception of a collision risk during an avoidance manoeuvre (Lee, 1976; Bootsma and Craig, 2003). The 2.5 s threshold appears to be consistent with the recommendations that were made by the U.S. Department of Transportation (NHTSA, 2013). Indeed, the minimum warning thresholds recommended in the test protocols for collision warning systems are 2.1, 2.4, and 2.0 s when the vehicle respectively approaches a fixed, decelerating or low-speed obstacle. During an avoidance manoeuvre, when the vehicle passed a pedestrian, the lateral offset also influenced risk perception. The closer the vehicle was to the pedestrian, the greater the subjective risk became. However, the results demonstrate that the offset had a smaller effect on the SA than the TTC did and may not have affected the SCR. Hence, subjective risk perception has evolved in the same way as EDA on average but not necessarily with the same magnitude.

The subjective assessment of collision risk is influenced by vehicle dynamics. The non-linearity of the effects observed on the indicators reveals that risk perception does not result from the simple relationship between the probability of a hazard and its importance. Rather, the results evoke a threshold effect as suggested by Boer (2006). Each passenger built up safety margins and would only perceive a risk when the vehicle approached a pedestrian and violated these margins. The passenger's risk perception would, therefore, result from a continuous confrontation between the vehicle's trajectory and their safety margins. These findings are compatible with the concept of the "safe field of travel" that was introduced by Gibson and Crooks (1938), according to which an individual represents a dynamic area in which their vehicle can navigate safely. In comparison with the experiment conducted by Ferrier-Barbut et al. (2018), who used a virtual-reality helmet to test the impact of proximity between pedestrians and a vehicle, this experiment utilized a driving simulator that lacked a physical vehicle cab, which may have made it difficult to estimate the vehicle's width and its lateral distance from object (Mecheri and Lobjois, 2018). Although Walker et al. (2019b) demonstrated that medium-level driving simulators remain appropriate for the study of risk perception, the lack of a physical cab simulator and the absence of real danger may have limited the participants' abilities to gauge their proximity to the pedestrians.

The participants' physiological responses to approaching pedestrians reflect the activation of the sympathetic nervous system that operates parallel to subjective evaluation. As Choi et al. (2019) stated, the sympathetic nervous system can only react to a certain level of danger, and this can cause variations in certain physiological variables. The analysis of the EDA in this experiment confirmed this idea by revealing an increase in indicators mostly during high subjective risk assessment.

This study investigated how passengers perceive risk in autonomous vehicles that are navigating areas that include pedestrians. The aim was to better understand the fundamental mechanisms of risk perception in the particular case of shared spaces. Understanding how vehicle-environment dynamics influence the perception of a vehicle's passengers and pedestrians could help researchers create and implement motion algorithms that are compatible with the safety margins of all agents in a system design approach. This could also condition the acceptability of autonomous vehicles. This study presents preliminary results regarding this topic. However, many other parameters of vehicle-environment dynamics must be studied to progress. The question regarding the variables that were chosen to evaluate the passenger's feelings based on subjective evaluations or physiological measurements is essential. These results demonstrate that the measurements of the two types of indicators are not independent but are instead complementary.

6. CONCLUSION

This study highlighted the relevance of declarative and physiological measures of the real-time analysis of risks perceived by those in an autonomous vehicle. The results obtained are consistent with the literature concerning the effects of the manipulated variables. The value of the TTC at the beginning of a pedestrian avoidance manoeuvre and the lateral distance left between the vehicle and the pedestrian do affect the subjective risk. This study has demonstrated that physiological and subjective indicators are not independent but do not always lead to the same results, which supports the proposition made by Février et al. (2011). They stated that declarative and physiological measures are not redundant but complementary. This experience has demonstrated that one must be careful to not make a universal conclusion based on a single indicator in studies on risk perception. Subjective evaluations (risk as analysis) may be more sensitive to low-risk situations than physiological responses (risk as feeling) in particular. This work and its conclusions would benefit from being replicated in more realistic and complex environments. The safety margins that were tested in the driving simulator may not fully match those tolerated in a real autonomous vehicle. In a simulator or vehicle, the relationship between the two risk perception systems could be evaluated in a more complex and realistically modeled space, for example by varying pedestrian behaviors. As the environment becomes more complex, new risk perception factors (in addition to vehicle-environment dynamics) could be revealed. For example, passenger perception could be affected by unpredictability or lack of understanding of the autonomous vehicle state. The modeling approach we adopted could also be implemented to assess differences in risk perception between active drivers and passengers. Indeed, Basu et al. (2017) have shown that passengers prefer a more defensive driving style than when they are themselves in control of the vehicle. Finally, other works could be interested in the interaction modalities to be considered for the communication of information to the passenger-drivers so that they feel more secure (Bengler et al., 2020).

DATA AVAILABILITY STATEMENT

The dataset and the algorithm used in this study are available as a R package that can be downloaded from: https://gitlab.univ-nantes.fr/petit-j-2/bnscore.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the non-interventional research ethics committee of Nantes University (CERNI, num. 10032021). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JP, CC, and FM designed the study, contributed to the data modeling and statistical analysis, and wrote the manuscript. JP conducted the experiment. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the French National Research Agency (Agence Nationale de la Recherche, HIANIC project, Grant no. ANR-17-CE22-0010). The authors thank the ELICC Doctoral School (n°603) for its contribution to the linguistic correction of the article.

ACKNOWLEDGMENTS

The authors are thankful to the participants for their voluntary participation to this experiment.

REFERENCES

Basu, C., Yang, Q., Hungerman, D., Singhal, M., and Dragan, A. D. (2017). "Do you want your autonomous car to drive like you?," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, HRI' 17* (New York, NY: ACM), 417–425.

Benedek, M., and Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *J. Neurosci. Methods* 190, 80–91. doi: 10.1016/j.jneumeth.2010.04.028

Bengler, K., Rettenmaier, M., Fritz, N., and Feierle, A. (2020). From HMI to HMIs: towards an HMI framework for automated driving. *Information* 11:61. doi: 10.3390/info11020061

Binder, D. A. (1978). Bayesian cluster analysis. *Biometrika* 65, 31–38. doi: 10.1093/biomet/65.1.31

Boer, E. R. (2006). "Driving on the tail of catastrophe: towards an understanding of the risk managing interaction between driver, vehicle, automation, traffic, and infrastructure," in *SICE-ICASE, 2006, International Joint Conference* (Busan), 33–40.

Bootsma, R. J., and Craig, C. M. (2003). Information used in detecting upcoming collision. *Perception* 32, 525–544. doi: 10.1068/p3433

Bootsma, R. J., Fayt, V., Zaal, F. T. J. M., and Laurent, M. (1997). On the information-based regulation of movement: what Wann (1996) may want to consider. *J. Exp. Psychol.* 23, 1282–1289.

Boucsein, W. (2012). *Electrodermal Activity*. Springer Science & Business Media. doi: 10.1007/978-1-4614-1126-0

Braithwaite, J. J., Watson, D. G., Jones, R., and Rowe, M. (2013). *A Guide for Analyzing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments*. Technical Report, 2nd version: Selective Attention & Awareness Laboratory (SAAL) Behavioural Brain Sciences Centre, University of Birmingham. Available online at: https://www.birmingham.ac.uk/documents/college-les/psych/saal/guide-electrodermal-activity.pdf

Bresson, R., Saraydaryan, J., Dugdale, J., and Spalanzani, A. (2019). Socially compliant navigation in dense crowds.

Chen, R., Sherony, R., and Gabler, H. C. (2016). *Comparison of Time to Collision and Enhanced Time to Collision at Brake Application during Normal Driving*. SAE Technical Paper 2016-01-1448. Warrendale, PA: SAE International. doi: 10.4271/2016-01-1448

Choi, B., Jebelli, H., and Lee, S. (2019). Feasibility analysis of electrodermal activity (EDA) acquired from wearable sensors to assess construction workers' perceived risk. *Saf. Sci.* 115, 110–120. doi: 10.1016/j.ssci.2019.01.022

Daviaux, Y., Bonhomme, E., Ivers, H., de Sevin, É., Micoulaud-Franchi, J.-A., Bioulac, S., et al. (2020). Event-related electrodermal response to stress: results from a realistic driving simulator scenario. *Hum. Factors* 62, 138–151. doi: 10.1177/0018720819842779

Denis, J.-B., and Scutari, M. (2015). *R'eseaux Bayésiens avec R*. EDP Sciences, Les Ulis.

Distefano, N., Leonardi, S., Pulvirenti, G., Romano, R., Merat, N., Boer, E., et al. (2020). Physiological and driving behaviour changes associated to different road intersections. *Eur. Trans.* 77:4. doi: 10.48295/ET.2020.77.4

Droulers, O., Lajante, M., and Lacoste-Badie, S. (2013). Apport de la démarche neuroscientifique á la mesure des émotions : importation d'une nouvelle méthode de mesure de l'activité électrodermale. *Décis. Market.* 72, 87–101. doi: 10.7193/dm.072.87.101

Dunn, P. K., and Smyth, G. K. (2005). Series evaluation of Tweedie exponential dispersion model densities. *Stat. Comput*. 15, 267–280. doi: 10.1007/s11222-005-4070-y

Dunn, P. K., and Smyth, G. K. (2008). Evaluation of Tweedie exponential dispersion model densities by Fourier inversion. *Stat. Comput*. 18, 73–86. doi: 10.1007/s11222-007-9039-6

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol*. 59, 255–278. doi: 10.1146/annurev.psych.59.103006.093629

Evans, J. S. B. T., and Stanovich, K. E. (2013). Dual-process theories of higher cognition: advancing the debate. *Perspect. Psychol. Sci*. 8, 223–241. doi: 10.1177/1745691612460685

Faas, S. M., Mathis, L.-A., and Baumann, M. (2020). External HMI for self-driving vehicles: which information shall be displayed? *Transport. Res. Part F: Traffic Psychol. Behav*. 68, 171–186. doi: 10.1016/j.trf.2019.12.009

Ferrier-Barbut, E., Vaufreydaz, D., David, J.-A., Lussereau, J., and Spalanzani, A. (2018). "Personal space of autonomous car's passengers sitting in the driver's seat," in *2018 IEEE Intelligent Vehicles Symposium (IV)* (Suzhou), 2022–2029.

Février, F., Gauducheau, N., Jamet, É., Rouxel, G., and Salembier, P. (2011). La prise en compte des affects dans le domaine des interactions homme-machine : quels modéles, quelles méthodes, quels bénéfices? *Le Travail Hum*. 74, 183–201.

Findlay, F. (2017). How to get Great Electrodermal Activity (EDA) data analysis essentials.

Fowles, D. C., Christie, M. J., Edelberg, R., Grings, W. W., Lykken, D. T., and Venables, P. H. (1981). Publication recommendations for electrodermal measurements. *Psychophysiology* 18, 232–239.

Franzén, J. (2008). Bayesian cluster analysis. 161.

Gibson, J. J., and Crooks, L. E. (1938). A theoretical field-analysis of automobile-driving. *Am. J. Psychol*. 51:453.

Great Britain and Department of Transport (2011). *Shared Space*. Norwich: TSO.

Hamilton-Baillie, B. (2008). Towards shared space. *Urban Des. Int*. 13, 130–138. doi: 10.1057/udi.2008.13

Hankin, R. K. S. (2006). Additive integer partitions in R. *J. Stat. Softw*. 16, 1–3. doi: 10.18637/jss.v016.c01

Hankin, R. K. S., and West, L. G. (2007). Set partitions in R. *J. Stat. Softw*. 23, 1–12. doi: 10.18637/jss.v023.c02

Hasan, M. M., and Dunn, P. K. (2012). Understanding the effect of climatology on monthly rainfall amounts in Australia using Tweedie GLMs. *Int. J. Climatol*. 32, 1006–1017. doi: 10.1002/joc.2332

Herrero-Fernández, D., Parada-Fernández, P., Oliva-Macías, M., and Jorge, R. (2020). The influence of emotional state on risk perception in pedestrians: a psychophysiological approach. *Saf. Sci*. 130:104857. doi: 10.1016/j.ssci.2020.104857

Kabtoul, M., Spalanzani, A., and Martinet, P. (2020). "Towards proactive navigation: a pedestrian-vehicle cooperation based behavioral model," in *ICRA 2020 - IEEE International Conference on Robotics and Automation, IEEE International Conference on Robotics and Automation Proceedings* (Paris), 6958–6964.

Kaparias, I., Bell, M. G. H., Miri, A., Chan, C., and Mount, B. (2012). Analysing the perceptions of pedestrians and drivers to shared space. *Transp. Res. Part F* 15, 297–310. doi: 10.1016/j.trf.2012.02.001

Kass, R. E., and Raftery, A. E. (1995). Bayes factors. *J. Am. Stat. Assoc*. 90, 773–795. doi: 10.1080/01621459.1995.10476572

Kolekar, S., de Winter, J., and Abbink, D. (2020a). Human-like driving behaviour emerges from a risk-based driver model. *Nat. Commun*. 11, 1–13. doi: 10.1038/s41467-020-18353-4

Kolekar, S., de Winter, J., and Abbink, D. (2020b). Which parts of the road guide obstacle avoidance? Quantifying the driver's risk field. *Appl. Ergon*. 89:103196. doi: 10.1016/j.apergo.2020.103196

Kyriakidis, M., de Winter, J. C. F., Stanton, N., Bellet, T., van Arem, B., Brookhuis, K., et al. (2019). A human factors perspective on automated driving. *Theor. Issues Ergon. Sci*. 20, 223–249. doi: 10.1080/1463922X.2017.1293187

Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception* 5, 437–459. doi: 10.1068/p050437

Loewenstein, G. F., Weber, E. U., Hsee, C. K., and Welch, N. (2001). Risk as feelings. *Psychol. Bull*. 127, 267–286. doi: 10.1037/0033-2909.127.2.267

MATLAB (2018). *Version 9.5.0.944444 (R2018b)*. Natick, MA: The MathWorks, Inc.

Mecheri, S., and Lobjois, R. (2018). Steering control in a low-cost driving simulator: a case for the role of virtual vehicle cab. *Hum. Factors* 60, 719–734. doi: 10.1177/0018720818769253

Métayer, N., and Coeugnet, S. (2021). Improving the experience in the pedestrian's interaction with an autonomous vehicle: an ergonomic comparison of external HMI. *Appl. Ergon*. 96:103478. doi: 10.1016/j.apergo.2021.103478

Moody, S., and Melia, S. (2014). Shared space: research, policy and problems. *Proc. Instit. Civil Eng*. 167:6. doi: 10.1680/tran.12.00047

Morange-Majoux, F. (2017). "Les indices psycho-physiologiques pour appréhender le comportement," in *Manuel visuel de psychophysiology*. Dunod: Paris, p. 113–131.

Näätänen, R., and Summala, H. (1976). *Road-User Behaviour and Traffic Accidents*. Amsterdam; New York, NY: North-Holland Publishing Company.

NHTSA (2013). *Forward Collision Warning System Confirmation Test*. Technical report, U.S. Department of Transportation, Washington, DC.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociol. Methodol*. 25, 111–163.

Raftery, A. E. (1999). Bayes factors and BIC: comment on "a critique of the Bayesian information criterion for model selection". *Sociol. Methods Res*. 27, 411–427. doi: 10.1177/0049124199027003005

Reilhac, P., Millett, N., and Hottelart, K. (2016). "Shifting paradigms and conceptual frameworks for automated driving," in *Road Vehicle Automation 3*, Lecture Notes in Mobility, eds G. Meyer and S. Beiker (Cham: Springer International Publishing), 73–89. doi: 10.1007/978-3-319-40503-2_7

Rossner, P., and Bullinger, A. C. (2019). "Do you shift or not? Influence of trajectory behaviour on perceived safety during automated driving on rural roads," in *HCI in Mobility, Transport, and Automotive Systems*, Lecture Notes in Computer Science, ed H. Krömker (Cham: Springer International Publishing), 245–254. doi: 10.1007/978-3-030-22666-4_18

Rossner, P., and Bullinger, A. C. (2020a). "How do you want to be driven? Investigation of different highly-automated driving styles on a highway scenario," in *Advances in Human Factors of Transportation, Advances in Intelligent Systems and Computing*, ed N. Stanton (Cham: Springer International Publishing), 36–43.

Rossner, P., and Bullinger, A. C. (2020b). "I care who and where you are– influence of type, position and quantity of oncoming vehicles on perceived safety during automated driving on rural roads," in *HCI in Mobility, Transport, and Automotive Systems*. Driving Behavior, Urban and Smart Mobility, Lecture Notes in Computer Science, ed H. Krömker (Cham: Springer International Publishing), 61–71.

Rota, G.-C. (1964). The number of partitions of a set. *Am. Math. Month*. 71, 498–504. doi: 10.1080/00029890.1964.11992270

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat*. 6, 461–464. doi: 10.1214/aos/1176344136

Slovic, P., Finucane, M. L., Peters, E., and MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Anal*. 24, 311–322. doi: 10.1111/j.0272-4332.2004.00433.x

Society for Psychophysiological Research *Ad Hoc* Committee on Electrodermal Measures (2012). Publication recommendations for electrodermal measurements. *Psychophysiology* 49, 1017–1034. doi: 10.1111/j.1469-8986.2012.01384.x

Stevens, S. S. (2017). *Psychophysics: Introduction to Its Perceptual, Neural and Social Prospects*. London; New York, NY: Routledge. doi: 10.4324/9781315127675

Verberne, F. M. F., Ham, J., and Midden, C. J. H. (2012). Trust in smart systems: sharing driving goals and giving information to increase trustworthiness and acceptability of smart systems in cars. *Hum. Factors* 54, 799–810. doi: 10.1177/0018720812443825

Vogel, K. (2003). A comparison of headway and time to collision as safety indicators. *Accid. Anal. Prevent*. 35, 427–433. doi: 10.1016/S0001-4575(02)00022-2

Walker, F., Dey, D., Martens, M., Pfleging, B., Eggen, B., and Terken, J. (2019a). "Feeling-of-safety slider: measuring pedestrian willingness to cross roads in field interactions with vehicles," in *CHI EA '19: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow), 1–6. doi: 10.1145/3290607.3312880

Walker, F., Hauslbauer, A. L., Preciado, D., Martens, M. H., and Verwey, W. B. (2019b). Enhanced perception of risk in a driving simulator. *Int. J. Hum. Factors Model. Simul.* 7, 100–118.

Zhang, Y. (2013). Likelihood-based and Bayesian methods for Tweedie compound Poisson linear mixed models. *Stat. Comput*. 23, 743–757. doi: 10.1007/s11222-012-9343-7

Zhao, C., Jiang, S., Lei, Y., and Wang, C. (2020). A study on an anthropomorphic car-following strategy framework of the autonomous coach in mixed traffic flow. *IEEE Access* 8, 64653–64665.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

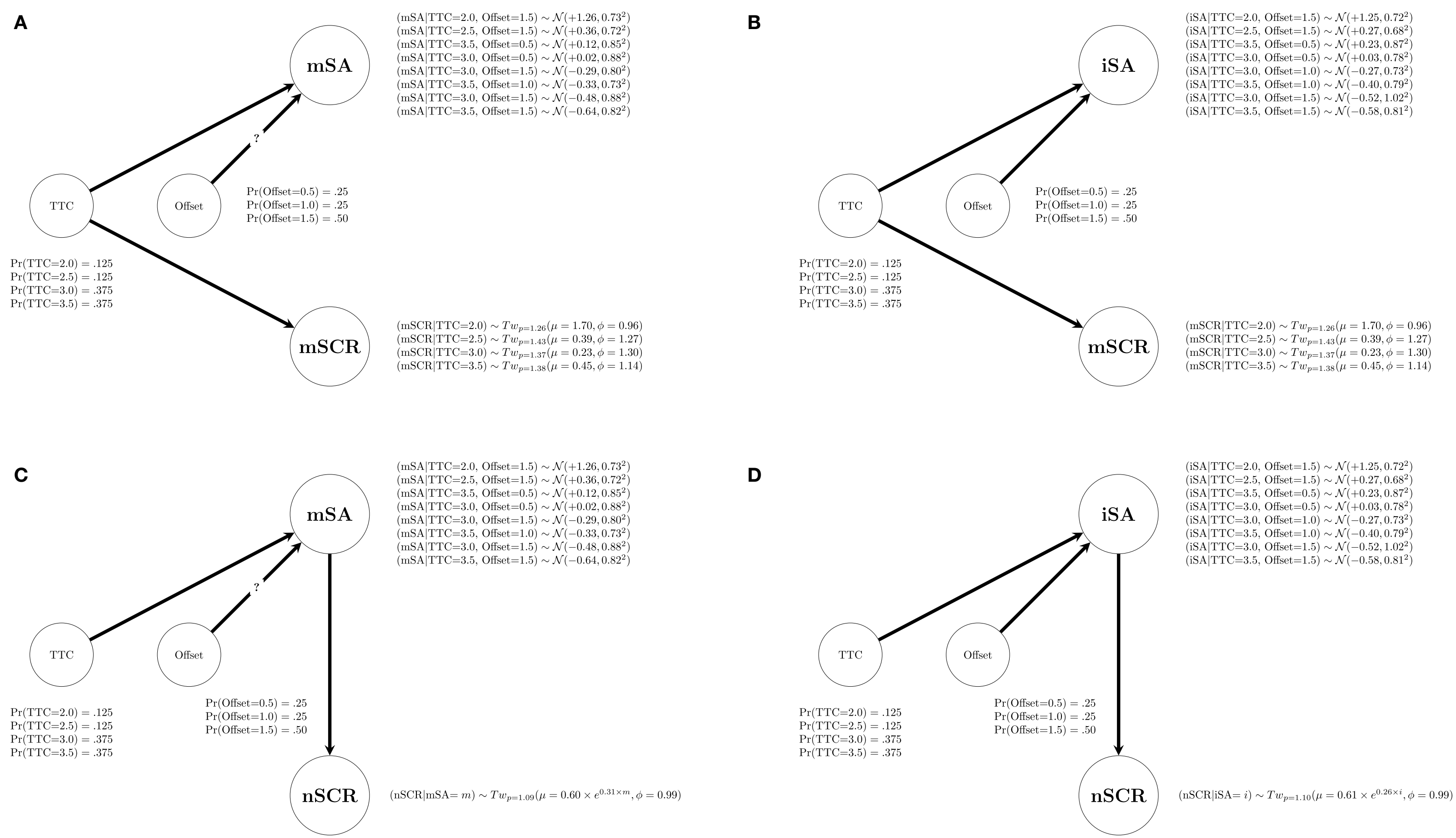


FIGURE A1 | The four best directed acyclic graphs with detailed distributions. **(A)** The relationships between the factors and the indicators mSA and mSCR. **(B)** The relationships between the factors and the indicators iSA and mSCR. **(C)** The relationships between the factors and the indicators mSA and nSCR. **(D)** The relationships between the factors and the indicators iSA and nSCR. The question mark on the relationship between node Offset and node mSA indicates that the degree of evidence for this relationship is not sufficient according to the BIC thresholds developed by Raftery (1995).

frontiers in Neuroergonomics

PERSPECTIVE
published: 29 November 2021
doi: 10.3389/fnrgo.2021.784827

Hybrid Systems to Boost EEG-Based Real-Time Action Decoding in Car Driving Scenarios

*Giovanni Vecchiato**

Institute of Neuroscience, National Research Council of Italy, Parma, Italy

The complexity of concurrent cerebral processes underlying driving makes such human behavior one of the most studied real-world activities in neuroergonomics. Several attempts have been made to decode, both offline and online, cerebral activity during car driving with the ultimate goal to develop brain-based systems for assistive devices. Electroencephalography (EEG) is the cornerstone of these studies providing the highest temporal resolution to track those cerebral processes underlying overt behavior. Particularly when investigating real-world scenarios as driving, EEG is constrained by factors such as robustness, comfortability, and high data variability affecting the decoding performance. Hence, additional peripheral signals can be combined with EEG for increasing replicability and the overall performance of the brain-based action decoder. In this regard, hybrid systems have been proposed for the detection of braking and steering actions in driving scenarios to improve the predictive power of the single neurophysiological measurement. These recent results represent a proof of concept of the level of technological maturity. They may pave the way for increasing the predictive power of peripheral signals, such as electroculogram (EOG) and electromyography (EMG), collected in real-world scenarios when informed by EEG measurements, even if collected only offline in standard laboratory settings. The promising usability of such hybrid systems should be further investigated in other domains of neuroergonomics.

Keywords: hybrid systems, action prediction, driving, EEG, EMG, EOG

OPEN ACCESS

Edited by:
Raphaëlle N. Roy,
Institut Supérieur de l'Aéronautique et de l'Espace (ISAE-SUPAERO), France

Reviewed by:
Angela Rose Harrivel,
National Aeronautics and Space Administration (NASA), United States
Caterina Cinel,
University of Essex, United Kingdom
Mathias Vukelić,
Fraunhofer Institute for Industrial Engineering, Germany

Correspondence:
Giovanni Vecchiato
giovanni.vecchiato@in.cnr.it

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: *28 September 2021*
Accepted: *08 November 2021*
Published: *29 November 2021*

Citation:
Vecchiato G (2021) Hybrid Systems to Boost EEG-Based Real-Time Action Decoding in Car Driving Scenarios. Front. Neuroergon. 2:784827. doi: 10.3389/fnrgo.2021.784827

INTRODUCTION

Today's human at work is asked to continuously interact with objects and the environment to perform a wide variety of tasks. In this regard, the research field of neuroergonomics aims to unravel the neural bases of those neurophysiological processes involved in the interaction between the user and a technical system during everyday life activities (Parasuraman, 2003; Dehais et al., 2020; Gramann et al., 2021).

Because of its complexity, one of the main real-world activities targeted by neuroergonomics is driving (Navarro et al., 2018). Studies demonstrated that driving behavior is the final result of simultaneous mental processes such as attention, decision-making, vigilance, motor, and cognitive control (Calhoun et al., 2002; Calhoun and Pearlson, 2012). Driving activity can cause drowsiness, fatigue, and an increase in workload, and it is one of the primary causes of death worldwide (Borghini et al., 2014). In order to assure road safety, it becomes fundamental to have a deep understanding of those mental processes underlying the interactions existing among the driver, the car, and the external environment to predict human behavior resulting in steering and braking actions. In recent years, driving scenarios have been enriched by technological advancements in designing autonomous cars (Badue et al., 2021). However, even if intelligent systems can execute actions on behalf of the driver, the correctness of these choices can only be evaluated once we understand those mechanisms underlying the driver's behavior in simulated and real traffic scenarios. In this regard, expertise could be an essential factor worth considering since evidence collected among professional and non-expert drivers suggest that the two populations share basic neurophysiological mechanisms, whereas the expertise subtending exceptional driving abilities may be associated with specific morphological and functional cerebral architecture changes (Bernardi et al., 2013, 2014).

For this reason, several pieces of research have been conducted to identify the neural basis of transportation and car driving. A recent meta-analysis presents a neuroergonomic framework according to which the neural bases of driving behavior are categorized into strategical (i.e., navigation), tactical (i.e., overtaking), and operational (steering and braking) tasks (Navarro et al., 2018).

In this context, developing efficient brain-based systems for the real-time decoding of brain processes underlying driver's behavior would be highly beneficial for the design of assistive devices.

This perspective provides a succinct overview of the literature about hybrid systems used for action prediction and the related limitations. Then, it presents the results related to car driving scenarios as proof of the level of technological maturity achieved in the last years. In this context, driving actions are predicted (i) exploiting secondary tasks eliciting cerebral activity related to a higher level of motor control and (ii) by measuring neural correlates of motor preparation as a marker of braking and steering actions and. This methodological approach could benefit additional ecological scenarios in neuroergonomics, such as telerehabilitation and occupational safety.

FROM EEG-BASED ACTION DECODERS TO HYBRID SYSTEMS

Electroencephalography (EEG) is one of the most used techniques for monitoring brain signals in operational environments. This measure provides the variation of electrical potentials on the scalp surface, generated by the summation of post-synaptic potentials within cortical layers (Biasiucci et al., 2019). Electroencephalography has the critical advantage of tracking brain dynamics with millisecond accuracy and is used in real-world scenarios for neuroimaging studies outside the lab. The evolution of technology allowed the removal of wires and produced wearable and long-lasting recording devices, enabling a wide range of experiments in real-world settings (Debener et al., 2012; Mihajlović et al., 2015; Mullen et al., 2015; Casson, 2019).

The integration of EEG-based action predictions into the control of an assistive technology device, such as a car, would have the great advantage of detecting, as early as possible, the movement preparation and execution, both in a laboratory and in more natural environmental settings. However, despite the noteworthy technological advancement of the last decade, there are still several issues that limit the utilization of the EEG for the real-time monitoring of actions in working environments. For instance, there is the need to improve the EEG hardware to obtain recordings more robust to artifacts and longer battery life and produce smaller devices to be socially accepted by everyone. Other psychophysiological and technological constraints make this prediction hard to achieve in real-life scenarios. Factors such as attention, memory load, fatigue, and competing cognitive processes (Gonçalves et al., 2006; Käthner et al., 2014; Calhoun and Adali, 2016), as well as user's individual characteristics such as lifestyle, gender, and age (Kasahara et al., 2015) influence brain dynamics producing significant intra- and inter-subject variability (Saha and Baumert, 2020; Saha et al., 2021). Common EEG artifacts generated by muscles and eye movements, impedance shifts, environmental noise are typically amplified in real-world scenarios, sensibly affecting the quality of EEG signals during real-time monitoring (Waard, 1996; Zander et al., 2017; Lohani et al., 2019). Also, wearing an EEG device for users within operational environments could be uncomfortable and lead to the corruption of the underlying brain processes. Although the technology provides researchers with high-impedance systems equipped with active shielded electrodes for mobile applications, these devices do not solve all the mentioned issues intrinsically characterizing all ecological environments. This low signal-to-noise ratio returned by raw EEG data requires the use of a range of conceptually very different and computationally expensive algorithms to extract significant temporal and frequency EEG features (Müller et al., 2004; Lotte et al., 2007, 2018; Krusienski et al., 2011; Bellotti et al., 2019). These algorithms often are demanding in terms of calibration because requiring large training sets and are not robust to real-life environmental noise affecting EEG recordings. Other issues relate to the high-dimensionality and non-stationarity of the EEG data, impacting the classification performance (Lotte et al., 2018). In addition, most of the classification methods used in the literature are applied for offline EEG analyses, thus requiring the improvement of this methodology for online applications to guarantee a computational efficiency for the real-time decoding of the brain activity. Hence, the computing hardware and software must warrant a sufficiently high performance and low latency to preserve the earliness of prediction (Wöhrle et al., 2017).

Hence, different physiological, behavioral, and technical data can be combined to improve the reliability of EEG-based predictions and their fully automated application for supporting the user in self-paced movements in critical environments. For example, the prediction of actions onset based on EEG analysis can be improved by the design of hybrid systems simultaneously monitoring additional peripheral signals, such as electroculogram (EOG) and electromyographic (EMG) data, depending on the context requirements (Kirchner et al., 2014). The hybrid concept was introduced in the field of the Brain Computer Interfaces (BCIs), exploiting advantages of different physiological signals and computational approaches to finally achieve specific goals better than a conventional EEG based system, such as improving the overall classification rate or reducing the rate of false positives (Pfurtscheller et al., 2010; Li et al., 2019). Hybrid systems should rely at least on one brain signal in the form of electrical, magnetic, or hemodynamic changes, and at the same time, they can incorporate peripheral or external signals to improve the whole system's performance. For instance, combinations of eye movement signals with neuronal signals usually are utilized for hybrid EEG–EOG BCIs (Usakli et al., 2009, 2010; Ma et al., 2015; Hong and Khan, 2017). Hence, the design of hybrid systems can improve the action prediction performance depending on the particular application.

Electroencephalography and EMG signals can be used to predict movements before the action onset reliably, showing that multimodal machine learning approaches can be potentially used to control an electronic device (Kirchner et al., 2014; Wöhrle et al., 2017). Unimodal EEG-based predictions can be achieved earlier with respect to unimodal EMG-based prediction, thus suggesting that EEG is more suitable for providing the user the feeling that a device delivers support on time without significant delay. Also, EEG analysis leads to more false positives than EMG due to the higher signal-to-noise ratio characterizing such neural data. In addition, which signals are relevant at which state of movement planning and execution have been systematically investigated with machine learning approaches to predict movement targets (Novak et al., 2013). This study reports that each sensing modality has its peculiarities. Electroencephalography is suitable for very early prediction or if the user cannot perform the movement. Electromyography and hand position are accurate after limb motion onset. Eye-tracking is accurate at motion onset, but it is not able to predict motion dynamics. Combining EEG and EOG results in higher accuracy than using a unimodal approach and is convenient since the two signals are often measured together. Augmenting EMG with eye-tracking allows predictions to be made earlier than with only EMG. However, this research field is not mature yet to make precise comparisons of performance and calibration times between machine learning approaches for unimodal and multimodal measurements.

Several challenges also characterize these hybrid systems. One of the significant issues in this research is identifying the best combinations of signals to reach the best prediction performance since the optimal combination could differ across users and experimental scenarios. Variables including system complexity, cost, user workload have to be evaluated when comparing hybrid systems with unimodal predictions. From the user's point of view, the complexity of hybrid systems is usually higher than that of conventional single modality recordings because they are required to wear multiple brain and body sensors. User acceptability is a crucial criterion that needs to be considered in designing and implementing such systems (Pfurtscheller et al., 2010; Li et al., 2019).

HYBRID SYSTEMS IN CAR DRIVING SCENARIOS

In the field of driving research, several studies addressed the issue of action detection and prediction based on the discrimination of different EEG features in simulated (Haufe et al., 2011; Gheorghe et al., 2013; Khaliliardali et al., 2015; Kim et al., 2015; Vecchiato et al., 2018, 2020, 2021) and real driving scenarios (Haufe et al., 2014; Zhang et al., 2015).

A few studies used secondary tasks to elicit neural features predicting steering during simulated and real car driving. In particular, the contingent negative variation (CNV) potential was generated by a go/no-go task to investigate the decoding of drive and brake events (Gheorghe et al., 2013; Khaliliardali et al., 2015). Results suggested that these actions can be discriminated around 320 ms before the movement with a classification performance of 0.77. In addition, Zhang et al. (2015) described an online event-related negativity (ERN) classifier to predict steering events, guided by a directional cue, both in a laboratory and real car driving scenarios. In both experimental conditions, they discriminated correct by error trials 480 and 700 ms after the directional cue. The classification performance is 0.70, but the computational timing cost is not reported, so the time interval between the directional cue and the classifier decision is unknown, and therefore whether it comes before or after the actual movement execution.

Other studies investigated the driver's action without using external cues with the advantage of limiting the additional driver's mental load. Haufe et al. (2011) explored pseudo-online emergency braking detection and evaluated that such a system in a simulation environment could eventually detect foot action around 130 ms before its onset. The possibility to decode self-generated actions detecting steering was also assessed, and in particular, whether the driver would perform a lane change in a simulated highway was predicted about 800 ms earlier the action onset with a true positive rate of 74.6% (Gheorghe et al., 2013). In addition, Vecchiato et al. (2018) identified an EEG independent component associated with the fronto-central electrodes exhibiting synchronization of theta EEG rhythm around 800 ms before the braking onset.

In line with the concept of hybrid systems, Kim et al. (2015) proposed a combination of EEG features in the time and frequency domains to distinguish three different kinds of stimulus-driven brake situations (i.e., sharp, soft, no brake). This study reported the highest EEG response-locked decoding performance at −480 and −420 ms distinguishing sharp and soft braking from no braking, respectively. It was harder to classify sharp and soft braking conditions with the same method (largest difference at −160 ms), which returned lower performance than a classifier based on EMG features. There were also significant results with hybrid decoding systems in a real car scenario where participants were asked to drive on a non-public test track (Haufe et al., 2014). They reported that a hybrid (EEG and behavioral features) classifier detected emergency braking even earlier than the laboratory setting (around 300 ms before the braking onset) (Haufe et al., 2011).

Moreover, in order to characterize the relative contribution of the EEG associated with the preparation of natural and self-initiated steering actions while driving to investigate its predictive power, the EEG related to *continuous steering* during the driving simulation was tested by means of canonical correlation analysis (CCA) and a linear lagged regression approach (LLR) to

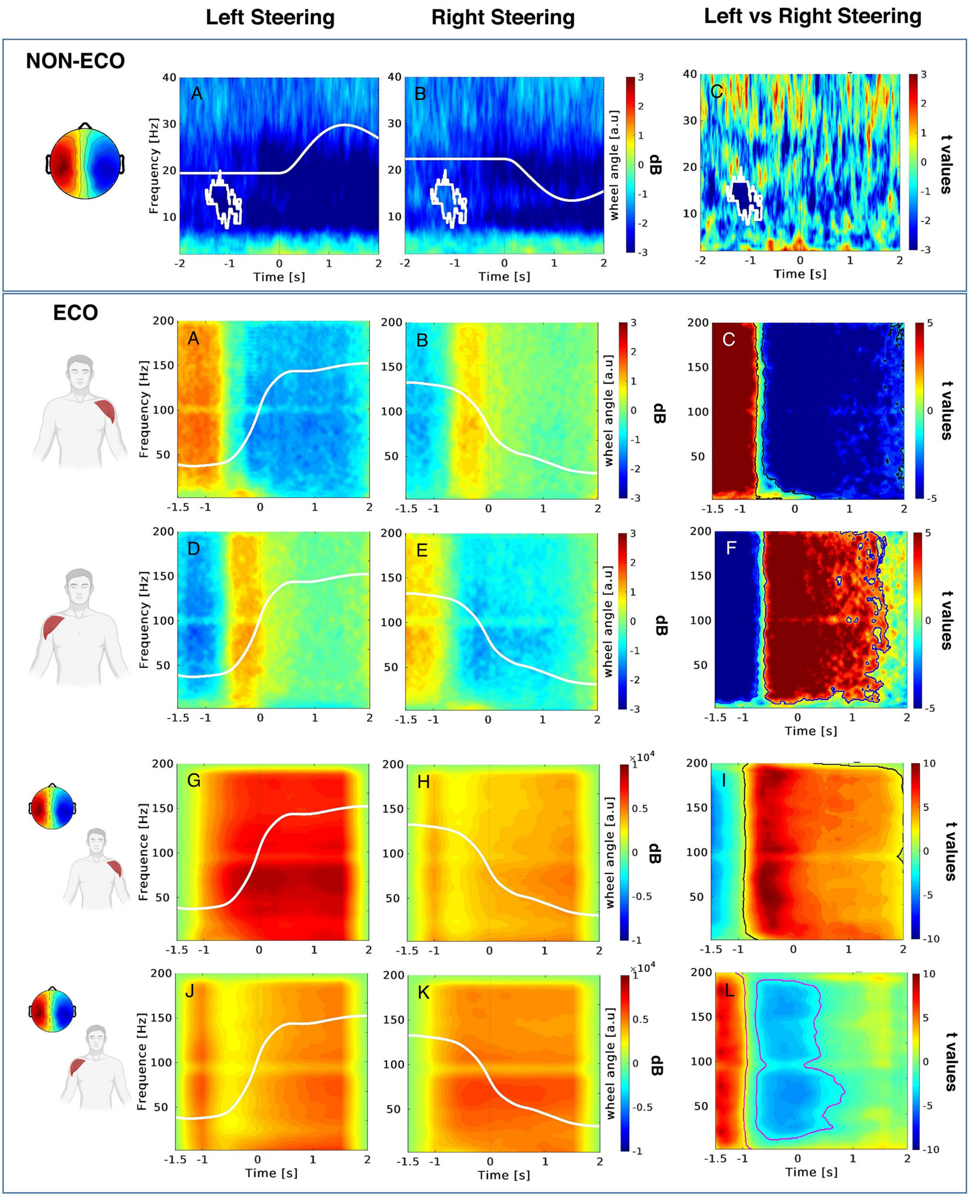


FIGURE 1 | NON-ECO frame. Time-frequency EEG patterns collected during left **(A)** and right **(B)** non-ecological steering, as well as their statistical comparison **(C)**. The topography in the left part of the picture shows the average scalp map related to the cluster of independent components. **ECO frame.** ERSP for the EMG signals *(Continued)*

FIGURE 1 | collected during the non-ecological steering task **(A–F)**, and cross-correlation results between EEG and EMG data **(G–L)**. The first (second) row (from the top) illustrates the EMG ERSP for the left (right) deltoid during left and right steering, as well as their statistical comparison. The third (fourth) row illustrates the EEG-EMG cross-correlation values for the left (right) deltoid during left and right steering, and the statistical comparison of the two conditions. White lines depict the left and right steering wheel angle profiles. Color bars indicate in blue (red) the decrease (increase) of EEG, EMG, and cross-correlation, as well as the statistical differences corresponding to the decrease (increase) of such activity during the left (right) steering. White and black masks delimit the statistically significant portion of the EEG, EMG, cross-correlation panels (adapted from Vecchiato et al., 2021).

identify the relative contribution of the EEG signals in steering anticipation (Di Liberto, accepted). Results showed that the combination of CCA-LLR analysis is valuable to disentangle the relative contribution of behavioral and electrophysiological components—within the EEG signals—for steering prediction in a continuous driving simulation task. This result demonstrates that brain-related EEG signals significantly improve the overall decoding performance, showing that the significant contribution in predicting steering comes from non-brain-related signals, such as ocular and muscular components.

Brain and muscular activities underlying steering behavior were also investigated with the final aim to increase the overall ecology of the experimental setting (Vecchiato et al., 2021). In particular, EEG feature predicting steering action and direction elicited by responding to traffic signs displayed on a computer screen was extracted and later exploited to increase the predictive power of the EMG collected in a more ecological steering task, such as a driving simulation. The desynchronization of the mu rhythm during the motor preparation of non-ecological steering cued by the traffic sign discriminated the muscular activity of the deltoids, thus anticipating subject steering behavior of 1.5 s. In addition, the increase of EMG activity of the deltoids anticipated the contralateral steering in both non-ecological and ecological steering tasks of 200 and 500 ms relative to the action onset, making it possible to discriminate such a driving behavior. Although these variations of EMG activity appear before the action onset allowing for possible online predictions, EEG data were used to increase the available time to perform such a calculation. The identified *non-ecological EEG* feature correlated with the *ecological EMG* activity of the deltoids, providing an improvement of the discrimination power of the steering side during driving simulation (**Figure 1**). These findings show an approach to increase the ecology of the experimental setting by limiting the invasiveness of the neurophysiological measurements using surface EMG sensors in the ecological scenario and combining neural data collected in the non-ecological one. This approach provides a way to monitor the user performance online through a simpler to acquire muscular correlate when compared to neural data, which could be recorded offline to increase the decoding system's performance without impacting the complexity of the ecological setting.

CONCLUSIONS

The coupling between EEG, EMG, and ocular signals is a valid mechanism for utilizing hybrid systems for the detection and online prediction of driving actions, exemplifying how it might be possible to complement information from behavioral, physiological, and external sources to control the level of assistance needed by the driver in that context (Chavarriaga et al., 2018). This methodology could pave the way for the utilization of hybrid systems based on neural signals—collected in standard laboratory settings and processed offline—having the role in improving the predictive power of peripheral signals—collected in more ecological settings and possibly processed online—correlated with the upcoming action execution.

The predictive power returned by coupling the EEG with peripheral signals demonstrated in car driving scenarios could be further investigated in larger sets of actions to extend the validity of this approach to other neuroergonomic areas. For instance, this methodology could foster the spread of mobile brain and body applications (Makeig et al., 2009; Gramann et al., 2011) and BCI paradigms (Douibi et al., 2021; Saha et al., 2021) onto several other contexts of our daily life. The capability to remotely monitor in an ecological way an individual's action would have a tremendous impact in the rehabilitation field (Nuara et al., 2021), with the possibility to verify the compliance and adherence to treatment relieving the patient and caregivers from a massive burden in terms of time and costs. Applications could also extend beyond the clinical realm, virtually to any fields where action surveillance would be valuable for preventing harmful consequences. It is the case, for instance, of the occupational safety of workers dealing in their routine with unsafe practices, for whom the use of this ecological methodology could reduce the likelihood of occupational injuries during the performance of high-risk motor tasks (Rizzolatti et al., 2021).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This study received funding from CAMLIN Limited. The funder was not involved in the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

The present study was supported by a research agreement between CAMLIN Limited and IN-CNR, and a research agreement between CAMLIN Limited and Toyota Motor Europe. The author would like to thank Maria Del Vecchio for the insightful comments she provided on the drafts of this manuscript.

REFERENCES

Badue, C., Guidolini, R., Carneiro, R. V., Azevedo, P., Cardoso, V. B., Forechi, A., et al. (2021). Self-driving cars: a survey. *Expert Syst. Appl.* 165, 113816. doi: 10.1016/j.eswa.2020.113816

Bellotti, A., Antopolskiy, S., Marchenkova, A., Colucciello, A., Avanzini, P., Vecchiato, G., et al. (2019). "Brain-based control of car infotainment," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)* (Bari), 2166–2173. doi: 10.1109/SMC.2019.8914448

Bernardi, G., Cecchetti, L., Handjaras, G., Sani, L., Gaglianese, A., Ceccarelli, R., et al. (2014). It's not all in your car: functional and structural correlates of exceptional driving skills in professional racers. *Front. Hum. Neurosci.* 8, 888. doi: 10.3389/fnhum.2014.00888

Bernardi, G., Ricciardi, E., Sani, L., Gaglianese, A., Papasogli, A., Ceccarelli, R., et al. (2013). How skill expertise shapes the brain functional architecture: an fMRI study of visuo-spatial and motor processing in professional racing-car and naïve drivers. *PLoS ONE* 8, e77764. doi: 10.1371/journal.pone.0077764

Biasiucci, A., Franceschiello, B., and Murray, M. M. (2019). Electroencephalography. *Curr. Biol.* 29, R80–R85. doi: 10.1016/j.cub.2018.11.052

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., and Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neurosci. Biobehav. Rev.* 44, 58–75. doi: 10.1016/j.neubiorev.2012.10.003

Calhoun, V. D., and Adali, T. (2016). Time-varying brain connectivity in fMRI data: whole-brain data-driven approaches for capturing and characterizing dynamic states. *IEEE Signal Process. Mag.* 33, 52–66. doi: 10.1109/MSP.2015.2478915

Calhoun, V. D., and Pearlson, G. D. (2012). A selective review of simulated driving studies: combining naturalistic and hybrid paradigms, analysis approaches, and future directions. *Neuroimage* 59, 25–35. doi: 10.1016/j.neuroimage.2011.06.037

Calhoun, V. D., Pekar, J. J., McGinty, V. B., Adali, T., Watson, T. D., and Pearlson, G. D. (2002). Different activation dynamics in multiple neural systems during simulated driving. *Hum. Brain Mapp.* 17, 141–142. doi: 10.1002/hbm.10066

Casson, A. J. (2019). Wearable EEG and beyond. *Biomed. Eng. Lett.* 9, 53–71. doi: 10.1007/s13534-018-00093-6

Chavarriaga, R., Ušćumlić, M., Zhang, H., Khaliliardali, Z., Aydarkhanov, R., Saeedi, S., et al. (2018). Decoding neural correlates of cognitive states to enhance driving experience. *IEEE Trans. Emerg. Top. Comput. Intell.* 2, 288–297. doi: 10.1109/TETCI.2018.2848289

Debener, S., Minow, F., Emkes, R., Gandras, K., and de Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk?: EEG to go. *Psychophysiology* 49, 1617–1621. doi: 10.1111/j.1469-8986.2012.01471.x

Dehais, F., Karwowski, W., and Ayaz, H. (2020). Brain at work and in everyday life as the next frontier: grand field challenges for neuroergonomics. *Front. Neuroergon.* 1, 583733. doi: 10.3389/fnrgo.2020.583733

Di Liberto, G., Barsotti, M., Vecchiato, G., Ambeck-Madsen, J., Del Vecchio, M., Avanzini, P., et al. (accepted). Robust anticipation of continuous steering actions from electroencephalographic data during simulated driving. *Sci. Rep.*

Douibi, K., Le Bars, S., Lemontey, A., Nag, L., Balp, R., and Breda, G. (2021). Toward EEG-based BCI applications for industry 4.0: challenges and possible applications. *Front. Hum. Neurosci.* 15, 705064. doi: 10.3389/fnhum.2021.705064

Gheorghe, L., Chavarriaga, R., and Millán, J. D. R. (2013). "Steering timing prediction in a driving simulator task," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Piscataway, NJ), 6913–6916. doi: 10.1109/EMBC.2013.6611147

Gonçalves, S. I., de Munck, J. C., Pouwels, P. J. W., Schoonhoven, R., Kuijer, J. P. A., Maurits, N. M., et al. (2006). Correlating the alpha rhythm to BOLD using simultaneous EEG/fMRI: inter-subject variability. *Neuroimage* 30, 203–213. doi: 10.1016/j.neuroimage.2005.09.062

Gramann, K., Gwin, J. T., Ferris, D. P., Oie, K., Jung, T.-P., Lin, C.-T., et al. (2011). Cognition in action: imaging brain/body dynamics in mobile humans. *Rev. Neurosci.* 22, 593–608. doi: 10.1515/RNS.2011.047

Gramann, K., McKendrick, R., Baldwin, C., Roy, R. N., Jeunet, C., Mehta, R. K., et al. (2021). Grand field challenges for cognitive neuroergonomics in the coming decade. *Front. Neuroergon.* 2, 643969. doi: 10.3389/fnrgo.2021.643969

Haufe, S., Kim, J.-W., Kim, I.-H., Sonnleitner, A., Schrauf, M., Curio, G., et al. (2014). Electrophysiology-based detection of emergency braking intention in real-world driving. *J. Neural Eng.* 11, 056011. doi: 10.1088/1741-2560/11/5/056011

Haufe, S., Treder, M. S., Gugler, M. F., Sagebaum, M., Curio, G., and Blankertz, B. (2011). EEG potentials predict upcoming emergency brakings during simulated driving. *J. Neural Eng.* 8, 056001. doi: 10.1088/1741-2560/8/5/056001

Hong, K.-S., and Khan, M. J. (2017). Hybrid brain-computer interface techniques for improved classification accuracy and increased number of commands: a review. *Front. Neurorobot.* 11, 35. doi: 10.3389/fnbot.2017.00035

Kasahara, K., DaSalla, C. S., Honda, M., and Hanakawa, T. (2015). Neuroanatomical correlates of brain-computer interface performance. *NeuroImage* 110, 95–100. doi: 10.1016/j.neuroimage.2015.01.055

Käthner, I., Wriessnegger, S. C., Müller-Putz, G. R., Kübler, A., and Halder, S. (2014). Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface. *Biol. Psychol.* 102, 118–129. doi: 10.1016/j.biopsycho.2014.07.014

Khaliliardali, Z., Chavarriaga, R., Gheorghe, L. A., and Millán, J., del R. (2015). Action prediction based on anticipatory brain potentials during simulated driving. *J. Neural Eng.* 12, 066006. doi: 10.1088/1741-2560/12/6/066006

Kim, I.-H., Kim, J.-W., Haufe, S., and Lee, S.-W. (2015). Detection of braking intention in diverse situations during simulated driving based on EEG feature combination. *J. Neural Eng.* 12, 016001. doi: 10.1088/1741-2560/12/1/016001

Kirchner, E. A., Tabie, M., and Seeland, A. (2014). Multimodal movement prediction - towards an individual assistance of patients. *PLoS ONE* 9, e85060. doi: 10.1371/journal.pone.0085060

Krusienski, D. J., Grosse-Wentrup, M., Galán, F., Coyle, D., Miller, K. J., Forney, E., et al. (2011). Critical issues in state-of-the-art brain-computer interface signal processing. *J. Neural Eng.* 8, 025002. doi: 10.1088/1741-2560/8/2/025002

Li, Z., Zhang, S., and Pan, J. (2019). Advances in hybrid brain-computer interfaces: principles, design, and applications. *Comput. Intell. Neurosci.* 2019, e3807670. doi: 10.1155/2019/3807670

Lohani, M., Payne, B. R., and Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Front. Hum. Neurosci.* 13, 57. doi: 10.3389/fnhum.2019.00057

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., et al. (2018). A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *J. Neural Eng.* 15, 031005. doi: 10.1088/1741-2552/aab2f2

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., and Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain-computer interfaces. *J. Neural Eng.* 4, R1–R13. doi: 10.1088/1741-2560/4/2/R01

Ma, J., Zhang, Y., Cichocki, A., and Matsuno, F. (2015). A novel EOG/EEG Hybrid human–machine interface adopting eye movements and ERPs: application to Robot Control. *IEEE Trans. Biomed. Eng.* 62, 876–889. doi: 10.1109/TBME.2014.2369483

Makeig, S., Gramann, K., Jung, T.-P., Sejnowski, T. J., and Poizner, H. (2009). Linking brain, mind and behavior. *Int. J. Psychophysiol.* 73, 95–100. doi: 10.1016/j.ijpsycho.2008.11.008

Mihajlović, V., Grundlehner, B., Vullers, R., and Penders, J. (2015). Wearable, wireless EEG solutions in daily life applications: what are we missing? *IEEE J. Biomed. Health Inform.* 19, 6–21. doi: 10.1109/JBHI.2014.2328317

Mullen, T. R., Kothe, C. A. E., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., et al. (2015). Real-time neuroimaging and cognitive monitoring using wearable dry EEG. *IEEE Trans. Biomed. Eng.* 62, 2553–2567. doi: 10.1109/TBME.2015.2481482

Müller, K.-R., Krauledat, M., Dornhege, G., Curio, G., and Blankertz, B. (2004). Machine learning techniques for brain-computer interfaces. *Biomed. Eng.* 49, 11–22. doi: 10.13109/9783666351419.11

Navarro, J., Reynaud, E., and Osiurak, F. (2018). Neuroergonomics of car driving: a critical meta-analysis of neuroimaging data on the human brain behind the wheel. *Neurosci. Biobehav. Rev.* 95, 464–479. doi: 10.1016/j.neubiorev.2018.10.016

Novak, D., Omlin, X., Leins-Hess, R., and Riener, R. (2013). Predicting targets of human reaching motions using different sensing technologies. *IEEE Trans. Biomed. Eng.* 60, 2645–2654. doi: 10.1109/TBME.2013.2262455

Nuara, A., Fabbri-Destro, M., Scalona, E., Lenzi, S. E., Rizzolatti, G., and Avanzini, P. (2021). Telerehabilitation in response to constrained physical distance: an opportunity to rethink neurorehabilitative routines. *J. Neurol.* 2021, 1–12. doi: 10.1007/s00415-021-10397-w

Parasuraman, R. (2003). Neuroergonomics: research and practice. *Theor. Issues Ergon. Sci.* 4, 5–20. doi: 10.1080/14639220210199753

Pfurtscheller, G., Allison, B. Z., Brunner, C., Bauernfeind, G., Solis-Escalante, T., Scherer, R., et al. (2010). The hybrid BCI. *Front. Neurosci.* 4, 30. doi: 10.3389/fnpro.2010.00003

Rizzolatti, G., Fabbri-Destro, M., Nuara, A., Gatti, R., and Avanzini, P. (2021). The role of mirror mechanism in the recovery, maintenance, and acquisition of motor abilities. *Neurosci. Biobehav. Rev.* 127, 404–423. doi: 10.1016/j.neubiorev.2021.04.024

Saha, S., and Baumert, M. (2020). Intra- and inter-subject variability in EEG-based sensorimotor brain computer interface: a review. *Front. Comput. Neurosci.* 13, 87. doi: 10.3389/fncom.2019.00087

Saha, S., Mamun, K. A., Ahmed, K., Mostafa, R., Naik, G. R., Darvishi, S., et al. (2021). Progress in brain computer interface: challenges and opportunities. *Front. Syst. Neurosci.* 15, 578875. doi: 10.3389/fnsys.2021.578875

Usakli, A. B., Gurkan, S., Aloise, F., Vecchiato, G., and Babiloni, F. (2009). "A hybrid platform based on EOG and EEG signals to restore communication for patients afflicted with progressive motor neuron diseases," in *Conference proceedings. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference 2009* (Minneapolis, MN), 543–546. doi: 10.1109/IEMBS.2009.5333742

Usakli, A. B., Gurkan, S., Aloise, F., Vecchiato, G., and Babiloni, F. (2010). On the use of electrooculogram for efficient human computer interfaces. *Comput. Intell. Neurosci.* 2010, 135629. doi: 10.1155/2010/135629

Vecchiato, G., Del Vecchio, M., Antopolskiy, S., Bellotti, A., Colucciello, A., Marchenkova, A., et al. (2020). "Alpha and beta EEG Desynchronizations anticipate steering actions in a driving simulation experiment," in *Intelligent Human Systems Integration 2020. Advances in Intelligent Systems and Computing*, eds T. Ahram, W. Karwowski, A. Vergnano, F. Leali, and R. Taiar (Cham: Springer International Publishing), 260–265. doi: 10.1007/978-3-030-39512-4_41

Vecchiato, G., Vecchio, M. D., Ambeck-Madsen, J., Ascari, L., and Avanzini, P. (2021). Hybrid EEG-EMG system to detect steering actions in car driving settings. *BioRxiv preprint*. doi: 10.1101/2021.09.16.460615

Vecchiato, G., Vecchio, M. D., Ascari, L., Antopolskiy, S., Deon, F., Kubin, L., et al. (2018). Electroencephalographic time-frequency patterns of braking and acceleration movement preparation in car driving simulation. *Brain Res.* 1716, 16–26. doi: 10.1016/j.brainres.2018.09.004

Waard, D. D. (1996). *The Measurement of Drivers' Mental Workload*. Available online at: https://www.semanticscholar.org/paper/The-measurement-of-drivers%27-mental-workload-Waard/a4e683f49e357cc5b21107285013e9bd470aefe2 (accessed September 7, 2021).

Wöhrle, H., Tabie, M., Kim, S. K., Kirchner, F., and Kirchner, E. A. (2017). A hybrid FPGA-Based system for EEG- and EMG-based online movement prediction. *Sensors* 17, 1552. doi: 10.3390/s17071552

Zander, T. O., Andreessen, L. M., Berg, A., Bleuel, M., Pawlitzki, J., Zawallich, L., et al. (2017). Evaluation of a dry EEG system for application of passive brain-computer interfaces in autonomous driving. *Front. Hum. Neurosci.* 11, 78. doi: 10.3389/fnhum.2017.00078

Zhang, H., Chavarriaga, R., Khaliliardali, Z., Gheorghe, L., Iturrate, I., Millán, J., et al. (2015). EEG-based decoding of error-related brain activity in a real-world driving task. *J. Neural Eng.* 12, 066028. doi: 10.1088/1741-2560/12/6/066028

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroergonomics

PERSPECTIVE
published: 08 December 2021
doi: 10.3389/fnrgo.2021.778043

Eye Tracking in Driver Attention Research—How Gaze Data Interpretations Influence What We Learn

Christer Ahlström[1,2†], Katja Kircher[1†], Marcus Nyström[3] and Benjamin Wolfe[4]*

[1] Swedish National Road and Transport Research Institute (VTI), Linköping, Sweden, [2] Department of Biomedical Engineering, Linköping University, Linköping, Sweden, [3] Lund University Humanities Lab, Lund, Sweden, [4] Department of Psychology, University of Toronto Mississauga, Mississauga, ON, Canada

Edited by:
Carryl L. Baldwin,
Wichita State University, United States

Reviewed by:
Daniel M. Roberts,
Proactive Life, Inc., United States
Sophie Lemonnier,
EA7312 Laboratoire de Psychologie Ergonomique et Sociale pour l'Expérience Utilisateurs (PERSEUS), France

***Correspondence:**
Marcus Nyström
marcus.nystrom@humlab.lu.se

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: 16 September 2021
Accepted: 01 November 2021
Published: 08 December 2021

Citation:
Ahlström C, Kircher K, Nyström M and Wolfe B (2021) Eye Tracking in Driver Attention Research—How Gaze Data Interpretations Influence What We Learn. Front. Neuroergon. 2:778043. doi: 10.3389/fnrgo.2021.778043

Eye tracking (ET) has been used extensively in driver attention research. Amongst other findings, ET data have increased our knowledge about what drivers look at in different traffic environments and how they distribute their glances when interacting with non-driving related tasks. Eye tracking is also the go-to method when determining driver distraction via glance target classification. At the same time, eye trackers are limited in the sense that they can only objectively measure the gaze direction. To learn more about why drivers look where they do, what information they acquire foveally and peripherally, how the road environment and traffic situation affect their behavior, and how their own expertise influences their actions, it is necessary to go beyond counting the targets that the driver foveates. In this perspective paper, we suggest a glance analysis approach that classifies glances based on their *purpose*. The main idea is to consider not only the intention behind each glance, but to also account for what is relevant in the surrounding scene, regardless of whether the driver has looked there or not. In essence, the old approaches, unaware as they are of the larger context or motivation behind eye movements, have taken us as far as they can. We propose this more integrative approach to gain a better understanding of the complexity of drivers' informational needs and how they satisfy them in the moment.

Keywords: eye tracking (ET), driving (veh), distraction and inattention, purpose-based analysis, coding scheme, context, relevance

INTRODUCTION

A video with an overlaid fixation cross that shows where the driver's gaze is focused relative to the scenery is a powerful visualization. From such data, it is possible to derive objective and quantitative results like gaze direction, dwell time, and glance frequency to objects and locations. In driver attention research, eye movement analysis has been used to learn more about gaze behavior associated with mobile phone use (Tivesten and Dozza, 2014), the distribution of eyes-off-road durations (Liang et al., 2012), where drivers look at the road to maintain a smooth travel path (Lappi et al., 2013), where drivers sample visual information when driving through intersections (Kircher and Ahlström, 2020), etc.

Despite everything that eye movement analysis has taught us about driver behavior, one should be aware of some fundamental limitations in using eye tracking (ET) to study driver attention and behavior. *First*, eye trackers only measure where and for how long we look in a certain direction or at a certain target. It is not a direct overt measure of visual attention (e.g., Deubel and Schneider, 1996), and information about the purpose of the glance or what information the brain cognitively processes during the glance can be very difficult to access (cf. Viviani, 1990). *Second*, there is no method to directly measure information acquisition via peripheral vision that works in real-world applications, even though research indicates that drivers are aware of much more than what is being foveated (Underwood et al., 2003). Wolfe et al. (2020) even argue that peripheral input provides much of the information the driver needs, both at a global level (the gist of the scene, acquired in parallel) and at a local level (providing information to guide search processes and eye movements more generally). *Third*, it has been shown that not all foveated information is processed (Simons, 2000; Mack, 2003). This is often referred to as looked but failed to see or inattentional blindness. *Finally*, eye movement data do not provide an easy way to determine whether the sampled information was relevant, necessary, and sufficient for the driver in the current situation (Kircher and Ahlström, 2018; Wolfe et al., 2020). Considering these limitations, it is clear that driver attention assessments cannot be based on single foveations, without also considering glance history and the present traffic situation.

An alternative to interpreting a driver's visual information sampling gaze by gaze, target by target, is to consider visual information acquisition in driving as a task where many different glance strategies can be equally appropriate. The basic idea is that an attentive driver has a "good enough" mental representation of the current situation, containing imperfect but adequate information about the surrounding scene (Summala, 2007; Hancock et al., 2009). As suggested by Wolfe et al. (2020), this mental representation is built from information acquired via a series of context-guided glances in combination with peripheral vision, using data from the attentive and pre-attentive stages of information acquisition, and possibly from other sources. The representation can only be sufficient if enough relevant information is included. We would need to know where and at what drivers look and for what reason (including what they see with peripheral vision), their intended travel path and other tasks they are doing, and preferably also their familiarity and experience with the given situation. The dilemma is that even with accurate ET, co-registered with a recording of the driver's environment, and an experimental design that controls for travel path and tasks, we still would not be able to measure (i) information sampled via peripheral vision and (ii) the top-down processes that are known to influence why and from where information is sampled (Kircher and Ahlström, 2018). Note that from a driver attention perspective, it is not even enough to investigate if the sampled information is relevant and if it has been sampled sufficiently, it is also necessary to check that no relevant information was missed. Still, when combined with additional data and an innovative data reduction approach, gaze data can still be an asset for monitoring driver attention.

In this perspective paper, we compare different approaches to encode and interpret ET data that has been used in the field of driver attention research. For each approach we discuss the data needed, the implicit or explicit definition of an attentive driver, the typical results that can be obtained, and the conclusions that are likely to be drawn (summarized in **Table 1**). In addition to classifying gaze data based on *direction* and on the foveated *target*, we also include an approach that classifies glances based on their *purpose*. In this paper, we argue that the purpose-based approach provides added value for understanding context-based driver attention.

DIRECTION-, TARGET-, AND PURPOSE-BASED EYE MOVEMENT INTERPRETATIONS

To understand the differences between the *direction*-, *target*-, and *purpose-based* approaches when studying driver attention with ET, we start with the illustration in **Figure 1**. A driver intending to continue straight ahead is approaching an intersection. At the same time, a bicyclist is leaving the intersection on the main road. The driver glances to the right, foveating the bicyclist.

Direction-Based Approach

In the direction-based approach, the gaze direction is registered, typically as forward, up, down, left, and right. This approach is typically used when the eye movements are recorded in a coordinate system that is fixed relative to a vehicle-mounted remote eye tracker. It is then easy to extract the gaze direction, without the need for a scene camera. In **Figure 1**, the direction-based approach would register a glance to the right.

The direction-based approach is often used to compute indicators like "eyes off road" or "percent road center" (PRC; Victor et al., 2005) and it can be employed in real-time with automated data encoding. A driver is considered attentive when directing a minimum percentage of glances within a sliding window to the "road center," which would be the relevant area. A drawback with this approach is that the relevant area is typically defined as "forward," regardless of where relevant information is positioned relative to the car. Data fusion makes it possible to define more elaborate relevant areas that can be coupled to the direction of the gaze. For example, if the eye tracker data are associated with a world model of the vehicle's cockpit, glances to relevant areas representing the speedometer, and the mirrors can be treated differently than other off-road glances (Ahlstrom et al., 2013). To some extent, situational circumstances can also be integrated via map data and proximity sensors, allowing automated adjustments of the relevant area(s), for example by taking road curvature (Ahlstrom et al., 2011) and intersections (Ahlström et al., 2021) into account. Data fusion with other data sources is still uncommon, and eye movements are to a large extent interpreted without situational information in the direction-based approach. It is thus unknown *what* the driver glances at and *why*.

TABLE 1 | Methodological aspects to consider when applying direction, target, and purpose-based approaches to eye tracking data.

Approach	Direction	Target	Purpose
Strategy	Identify the glance direction	Identify the foveated glance target	Identify the probable reason for the glance
Actual coding	To the right OR away from forward	Bicyclist	Checking for relevant traffic from right
External knowledge needed for classification	Coordinate system determining forward	View of outside world	View of outside world, traffic rules that apply, intended direction of travel
Coding method	Real-time automated coding is available	Manual or semi-automated	Manual
Typical result	Frequency and duration of eyes-off-road	Frequency and duration of glances toward (type of) target or area	Frequency and duration (or neglection) of target or area in context of relevance
Typical research questions	How much do drivers look in certain directions or away from the forward roadway?	How much do drivers look at various targets?	How often are relevant areas or targets neglected?

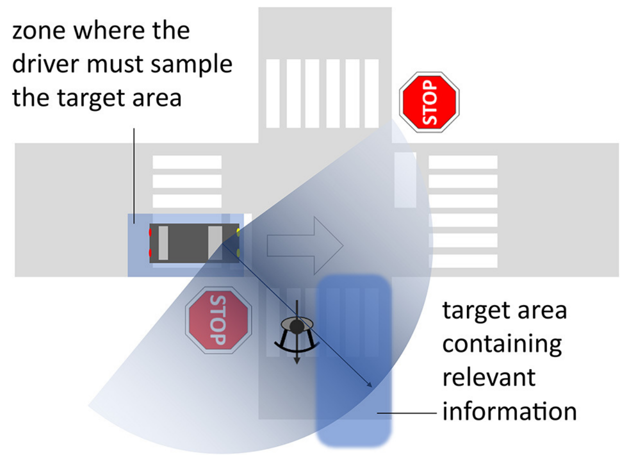


FIGURE 1 | Illustrative example showing a driver who is approaching an intersection with the intention to drive straight ahead. A bicyclist is leaving the intersection. While in the zone with a good view of the intersection, before entering it, the driver is looking right (thin arrow—foveal vision, shaded sector—peripheral vision), checking for traffic potentially present in the target area. A similar check for traffic from the left is required, too (but not illustrated in the figure).

Target-Based Approach

The *what*-question is typically answered by manual coding of scene videos with a gaze overlay, either from a remote or a head-mounted eye tracker. Solutions based on deep learning are also emerging where it is possible to automatically recognize objects in the videos and denote when the point-of-gaze intersects these objects (Panetta et al., 2020). To distinguish targets with similar XY-coordinates it may also be possible to use depth information from binocular eye trackers.

Glance *targets* are coded according to the target type, such as "bicyclist," "traffic sign," or "mobile phone" (Kircher and Ahlström, 2020). The glance to the right in **Figure 1** would be coded as "bicyclist." Target types that have a connection to traffic are often tacitly assumed to be relevant for driving, regardless of whether they contribute any relevant information in the current situation or not, like the bicyclist in **Figure 1** who is not relevant considering the driver's upcoming travel path. *Direction-* and *target-based* approaches commonly infer driver distraction when glances are directed away from forward or toward target types that are deemed irrelevant for driving (Halin et al., 2021). While widely accepted, these approaches often miss the important aspects of context, if relevant information is not foveated, and whether enough information is sampled with respect to the task at hand.

Purpose-Based Approach

For driver attention assessment, the *purpose-based approach* specifically defines which areas a driver must acquire information from to be considered attentive. This requires knowledge about the traffic rules that apply, which, in combination with the situation at hand, indicate where relevant information can be expected given the driver's intended maneuver. To assess the likely reason for a glance (or the absence of a glance), one must also consider the glance history, the infrastructure layout, other road users and traffic regulations. For example, a first glance down the crossing main road is likely meant to check for the presence of traffic. A follow-up glance in the same direction may help determining the available time gap for crossing the road. Here, the speed of the approaching road user may be more important than whether it is a car or a bicyclist. If all areas identified as relevant in the situation have been sampled timely and sufficiently, the driver will be considered attentive according to the purpose-based approach.

The theory of Minimum Required Attention (MiRA; Kircher and Ahlström, 2017) can be used as framework for an *a priori* definition of relevant areas. In **Figure 1**, one relevant area would be where traffic from the right can be expected ("target area"), regardless of whether traffic is present or not, and the purpose of the glance would thus be to check for traffic. Associated with the target area is a MiRA "zone," within which the driver must sample information from the target area. This zone is located on the driver's path and its shape is determined by situational circumstances like traffic regulations, line-of-sight, and intended direction of travel. This approach acknowledges that not only the presence but also the absence of other road users is relevant information.

The purpose-based approach explicitly includes the concept of spare capacity (cf. Kujala et al., 2021) by accepting glances to irrelevant areas/targets if all relevant targets are sampled sufficiently. So far, there is no straightforward method to determine when sampling is sufficient, and it appears as if foveal glances are not even necessary in all cases (Wolfe et al., 2017; Vater et al., 2020). Factors like presence, type, trajectory, and speed of other road users are likely to influence sufficiency.

Taking purpose into account leads to a rather different interpretation of the glance in **Figure 1**. Before crossing the intersection, the driver must check for traffic on the main road. The glance, especially if it is the first glance to the right in this location, is likely intended to check for traffic with right-of-way. With no such traffic present, the salient bicyclist happens to be foveated, even though the bicyclist is not relevant for the driver's upcoming maneuver. A purpose-based interpretation of the glance would be that the driver checked for traffic from the right as required, regardless of the actual target. To determine whether the driver was attentive in the given context, a glance checking for traffic from the left is required too, before the intersection is crossed.

DISCUSSION

Informative, useful, ET analyses rely on appropriate and reliable gaze data encodings and as we have discussed, these are tools that must be understood in a larger context. The automated data encodings that can be used in direction-based analyses have high objectivity, but they are not always appropriate, because they ignore where in relation to the environment the driver looked and why they looked where they did. For example, coding a glance as "eyes off road" when the driver's gaze is directed to the left (instead of forward) in an upcoming curve is incorrect, because it ignores this context. Opting for a target-based approach, asking what specific object the driver looked at, gives the impression of being more objective and accurate. After all, the driver's gaze either focused on a target or it did not, but the situation is not that simple. A driver's glance over their shoulder may end up being coded as a glance to the guardrail, because that is where foveation happened to occur, even though the intention was to check for overtaking traffic with peripheral vision, which renders the exact location of the fixation irrelevant in the process of acquiring the sought information. This clearly shows the dilemma of having to choose between an almost certainly wrong, but highly reliable coding of the fixated target, and a likely more correct purpose coding, which requires task knowledge and interpretation by the analyst. At least from research in sports there are indications, that in certain situations people fall back on purposely using peripheral vision to save energy and reduce suppression of visual input while the eyes are moving (see also Kredel et al., 2017; Vater et al., 2020).

To ensure reliability in a setting where the analyst's interpretations affect the results, it is important to use data encoding schemes that are well-founded in theoretical models and that suit the research question. In this paper, we use the MiRA theory (Kircher and Ahlström, 2017) to construct our model, although this is not the only possible approach. For example, the safety protocols suggested by Hirsch (1995) could be similarly useful. We do not argue that this approach is the one perfect solution to the problems we have pointed out in ET analyses, merely that it solves some of them. For example, the MiRA theory outlines how relevant areas can be defined, but it does not specify how drivers acquire information from these areas, if foveal vision is required, or if information acquisition via peripheral vision or other sensor modalities is enough. It is important to realize that the chosen theoretical model shapes the coding scheme and dictates what the analyst must infer from observed data. Both aspects have large consequences on the results. As with any new approach, effort must be made to ensure reliability and repeatability. Triangulation with other methods, as well as inter-rater reliability assessments, are good sanity checks for any approach with as many subjective elements as one which includes questions of motivation and reason. That said, being mindful of these limitations, a subjective purpose-based encoding can be more informative than an allegedly objective encoding of glance targets, and regardless of the approach chosen, *a priori* decisions must be made about the data coding scheme.

A key concern underlying our work here, which is unlikely to be alleviated in the near future, is the fact that eye trackers can only measure the gaze direction. They cannot measure information acquired via peripheral vision (Wolfe et al., 2020), spare visual capacity and acquisition of redundant information (Kujala et al., 2021), if fixated targets have been sampled sufficiently (Kircher and Ahlström, 2017), and what is known from past experience (Clark, 2015). In any model determining driver attention, merely knowing where a driver looked is neither sufficient nor adequate. Triangulating data, from multiple methods such as ET (including combinations of the direction-, target-, and purpose based approaches), driving behavior, think aloud (Ericsson and Simon, 1980), visual occlusion (Kujala et al., 2021), and event-related brain potentials (Hopstaken et al., 2016) with theoretical models of peripheral vision and neurocognitive function are likely to be necessary to attain a deeper understanding of driver attention (Kircher and Ahlström, 2018). As an example, by triangulating visual occlusion and ET results, it has been shown that glancing away from the forward roadway for driving purposes is not the same as glancing away for other purposes, and neither is necessarily equivalent to distraction (Kircher et al., 2019). This is, of course, not the only path that could lead to these conclusions, merely one among many.

On the whole, the data that eye trackers provide to driving researchers is immensely valuable, but like any other tool at the researcher's disposal, cannot be viewed as the one arbiter of truth. In this perspective paper, we have laid out ways in which ET data can both be used to better explain the complexities of driver behavior, and how particular ways in which they have been used can be misleading. Future ET research should consider the strengths and weaknesses we have detailed here, with particular attention to why drivers look where they do, what information they acquire foveally and peripherally, how the physical structure of the road environment dictates their behavior, and how their own expertise influences their acquisitive actions. The approach we advocate represents a significant shift in how ET data are used and understood, but it promises to provide key insights into what drivers need to know in a given situation and how they set about gaining the knowledge they require. In essence, the old approaches, unaware as they were of the larger context or motivation behind eye movements, have taken us as far as they can; we propose this complementary and more integrative approach to help researchers understand the complexity of drivers' informational needs and how they satisfy them in the moment.

AUTHOR CONTRIBUTIONS

The original ideas for the paper came from KK and CA. All authors contributed equally to the conception of the paper, development of the argument, and writing of the final version.

FUNDING

This work was partly funded by the Swedish Strategic Vehicle Research and Innovation Programme (VINNOVA FFI; Grant Number 2019-05834), by Stiftelsen Länsförsäkringsbolagens Forskningsfond, and by the Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN-2021-02730).

REFERENCES

Ahlström, C., Georgoulas, G., and Kircher, K. (2021). Towards a context-dependent multi-buffer driver distraction detection algorithm. *IEEE Trans. Intell. Transport. Syst.* 2021, 1–13. doi: 10.1109/TITS.2021.3060168

Ahlstrom, C., Kircher, K., and Kircher, A. (2013). A gaze-based driver distraction warning system and its effect on visual behavior. *IEEE Trans. Intell. Transport. Syst.* 14, 965–973. doi: 10.1109/TITS.2013.2247759

Ahlstrom, C., Victor, T., Wege, C., and Steinmetz, E. (2011). Processing of eye/head-tracking data in large-scale naturalistic driving data sets. *IEEE Trans. Intell. Transport. Syst.* 13, 553–564. doi: 10.1109/TITS.2011.2174786

Clark, A. (2015). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind.* New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780190217013.001.0001

Deubel, H., and Schneider, W. X. (1996). Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vision Res.* 36, 1827–1837. doi: 10.1016/0042-6989(95)00294-4

Ericsson, K. A., and Simon, H. A. (1980). Verbal reports as data. *Psychol. Rev.* 87, 215–251. doi: 10.1037/0033-295X.87.3.215

Halin, A., Verly, J. G., and Van Droogenbroeck, M. (2021). Survey and synthesis of state of the art in driver monitoring. *Sensors.* 21, 5558. doi: 10.3390/s21165558

Hancock, P. A., Mouloua, M., and Senders, J. W. (2009). "On the philosophical foundations of the distracted driver and driving distraction," in *Driver Distraction: Theory, Effects, and Mitigation,* eds M. A. Regan, J. D. Lee, and K. L. Young (Boca Raton, FL: CRC Press), 11–30. doi: 10.1201/9781420007497.pt2

Hirsch, P. (1995). "Proposed definitions of safe driving: an attempt to clear the road for more effective driver education," *Paper presented at the Canadian Multidisciplinary Road Safety Conference IX* (Montreal, QC).

Hopstaken, J. F., van der Linden, D., Bakker, A. B., Kompier, M. A., and Leung, Y. K. (2016). Shifts in attention during mental fatigue: evidence from subjective, behavioral, physiological, and eye-tracking data. *J. Exp. Psychol. Hum. Percept. Perform.* 42, 878. doi: 10.1037/xhp0000189

Kircher, K., and Ahlström, C. (2017). Minimum required attention: a human-centered approach to driver inattention. *Hum. Factors* 59, 471–484. doi: 10.1177/0018720816672756

Kircher, K., and Ahlström, C. (2018). Evaluation of methods for the assessment of attention while driving. *Accid. Anal. Prevent.* 114, 40–47. doi: 10.1016/j.aap.2017.03.013

Kircher, K., and Ahlström, C. (2020). Attentional requirements on cyclists and drivers in urban intersections. *Transport. Res. F Traff. Psychol. Behav.* 68, 105–117. doi: 10.1016/j.trf.2019.12.008

Kircher, K., Kujala, T., and Ahlström, C. (2019). On the difference between necessary and unnecessary glances away from the forward roadway: an occlusion study on the motorway. *Hum. Factors* 62, 1117–1131. doi: 10.1177/0018720819866946

Kredel, R., Vater, C., Klostermann, A., and Hossner, E.-J. (2017). Eye-tracking technology and the dynamics of natural gaze behavior in sports: a systematic review of 40 years of research. *Front. Psychol.* 8, 1845. doi: 10.3389/fpsyg.2017.01845

Kujala, T., Kircher, K., and Ahlström, C. (2021). A review of occlusion as a tool to assess attentional demand in driving. *Hum. Factors* 2021, 00187208211010953. doi: 10.1177/00187208211010953

Lappi, O., Lehtonen, E., Pekkanen, J., and Itkonen, T. (2013). Beyond the tangent point: gaze targets in naturalistic driving. *J. Vis.* 13, 11–11. doi: 10.1167/13.13.11

Liang, Y., Lee, J. D., and Yekhshatyan, L. (2012). How dangerous is looking away from the road? Algorithms predict crash risk from glance patterns in naturalistic driving. *Hum. Factors* 54, 1104–1116. doi: 10.1177/0018720812446965

Mack, A. (2003). Inattentional blindness: looking without seeing. *Curr. Dir. Psychol. Sci.* 12, 180–184. doi: 10.1111/1467-8721.01256

Panetta, K., Wan, Q., Rajeev, S., Kaszowska, A., Gardony, A. L., Naranjo, K., et al. (2020). ISeeColor: method for advanced visual analytics of eye tracking data. *IEEE Access* 8, 52278–52287. doi: 10.1109/ACCESS.2020.2980901

Simons, D. J. (2000). Attentional capture and inattentional blindness. *Trends Cogn. Sci.* 4, 147–155. doi: 10.1016/S1364-6613(00)01455-8

Summala, H. (2007). "Towards understanding motivational and emotional factors in driver behaviour: comfort through satisficing," in *Modelling Driver Behaviour in Automotive Environments: Critical Issues in Driver Interactions with Intelligent Transport Systems,* ed P. C. Cacciabue, (London: Springer), 189–207. doi: 10.1007/978-1-84628-618-6_11

Tivesten, E., and Dozza, M. (2014). Driving context and visual-manual phone tasks influence glance behavior in naturalistic driving. *Transport. Res. F Traff. Psychol. Behav.* 26, 258–272. doi: 10.1016/j.trf.2014.08.004

Underwood, G., Chapman, P., Berger, Z., and Crundall, D. (2003). Driving experience, attentional focusing, and the recall of recently inspected events. *Transport. Res. F Traff. Psychol. Behav.* 6, 289–304. doi: 10.1016/j.trf.2003.09.002

Vater, C., Williams, A. M., and Hossner, E.-J. (2020). What do we see out of the corner of our eye? The role of visual pivots and gaze anchors in sport. *Int. Rev. Sport Exerc. Psychol.* 13, 81–103. doi: 10.1080/1750984X.2019.1582082

Victor, T. W., Harbluk, J. L., and Engström, J. A. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transport. Res. F Traff. Psychol. Behav.* 8, 167–190. doi: 10.1016/j.trf.2005.04.014

Viviani, P. (1990). Eye movements in visual search: cognitive, perceptual and motor control aspects. *Rev. Oculomot. Res.* 4, 353–393.

Wolfe, B., Dobres, J., Rosenholtz, R., and Reimer, B. (2017). More than the useful field: considering peripheral vision in driving. *Appl. Ergon.* 65, 316–325. doi: 10.1016/j.apergo.2017.07.009

Wolfe, B., Sawyer, B. D., and Rosenholtz, R. (2020). Toward a theory of visual information acquisition in driving. *Hum. Factors* 2020, 0018720820939693. doi: 10.1177/0018720820939693

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroergonomics

SYSTEMATIC REVIEW
published: 14 December 2021
doi: 10.3389/fnrgo.2021.756473

Sensitivity of Physiological Measures of Acute Driver Stress: A Meta-Analytic Review

Laora Keraut ret[1,2], Stephanie Dabic[2] and Jordan Navarro[1,3]*

[1] Laboratoire d'Etude des Mécanismes Cognitifs, University Lyon 2, Lyon, France, [2] Valeo Interior Controls, Annemasse, France, [3] Institut Universitaire de France, Paris, France

Background: The link between driving performance impairment and driver stress is well-established. Identifying and understanding driver stress is therefore of major interest in terms of safety. Although many studies have examined various physiological measures to identify driver stress, none of these has as yet been definitively confirmed as offering definitive all-round validity in practice.

Aims: Based on the data available in the literature, our main goal was to provide a quantitative assessment of the sensitivity of the physiological measures used to identify driver stress. The secondary goal was to assess the influence of individual factors (i.e., characteristics of the driver) and ambient factors (i.e., characteristics of the context) on driver stress. Age and gender were investigated as individual factors. Ambient factors were considered through the experimental apparatus (real-road vs. driving simulator), automation driving (manual driving vs. fully autonomous driving) and stressor exposure duration (short vs. long-term).

Method: Nine meta-analyses were conducted to quantify the changes in each physiological measure during high-stress vs. low-stress driving. Meta-regressions and subgroup analyses were performed to assess the moderating effect of individual and ambient factors on driver stress.

Results: Changes in stress responses suggest that several measures are sensitive to levels of driver stress, including heart rate, R-R intervals (RRI) and pupil diameter. No influence of individual and ambient factors was observed for heart rate.

Applications and Perspective: These results provide an initial guide to researchers and practitioners when selecting physiological measures for quantifying driver stress. Based on the results, it is recommended that future research and practice use (*i*) multiple physiological measures, (*ii*) a triangulation-based methodology (combination of measurement modalities), and (*iii*) a multifactorial approach (analysis of the interaction of stressors and moderators).

Keywords: driver, stress, physiological, measures, sensitivity, individual, ambient, meta-analysis

Edited by:
Giovanni Vecchiato,
National Research Council of Italy, Italy

Reviewed by:
Giacinto Barresi,
Italian Institute of Technology (IIT), Italy
Mickael Causse,
National Higher School of Aeronautics and Space, France

****Correspondence:***
Laora Kerautret
laora.kerautret@gmail.com

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: *10 August 2021*
Accepted: *19 November 2021*
Published: *14 December 2021*

Citation:
Kerautret L, Dabic S and Navarro J (2021) Sensitivity of Physiological Measures of Acute Driver Stress: A Meta-Analytic Review. Front. Neuroergon. 2:756473. doi: 10.3389/fnrgo.2021.756473

INTRODUCTION

Identifying Driver Stress: A Safety and Comfort Challenge

Driving is a complex activity that takes place in a dynamic environment where safety critical situations abound. Therefore, many driving situations can lead the driver to experience stress, such as bad weather, low visibility, complex driver-environment interactions, and particular driving routes (Hill and Boyle, 2007; Rodrigues et al., 2015; Rastgoo et al., 2018). Although driver stress can be experienced as positive (i.e., eustress), the focus here is placed on its negative dimension (i.e., distress), which is more critical for well-being and road safety (Chung et al., 2019). Associated with negative emotions (e.g., anxiety, Kontogiannis, 2006, fear, Schmidt-Daffy, 2013, anger, Emo et al., 2016; Ooi et al., 2018; Gotardi et al., 2019) and the subjective feeling that the situation exceeds the individual's coping abilities (Selye, 1976), distress can lead to poor driving performances and risky behaviors (Matthews et al., 1998; Hancock and Desmond, 2001; Ge et al., 2014; Rendon-Velez et al., 2016). Given the causal relationship between distress and poor driving performance, finding measures that are sensitive to the level of stress is crucial if we are to gain a better understanding of this disturbed state and develop future remediation and support strategies.

Driver stress has often been identified on the basis of various subjective scales, including the Driver Stress Inventory (Matthews et al., 1997) and Driver Behavior Inventory (Gulian et al., 1989; Glendon et al., 1993). Although these scales have proven useful for capturing the multifaceted nature of driver stress, they may also be limited by individuals' inaccuracy in self-reporting stress levels. What is more, relationships with the neuroticism dimension have been shown to account for some of the inaccuracy of subjective stress ratings (McCrae, 1990; Espejo et al., 2011). Driver stress has also been inferred to a large extent from the analysis of driving behaviors, such as steering wheel motion, speed, acceleration, braking, overtaking, and lane keeping (Schießl, 2008; Rigas et al., 2012; Lanatà et al., 2014; Miller and Boyle, 2015; Rendon-Velez et al., 2016; Lee et al., 2017). Again, this method of identifying driver stress has some disadvantages. In addition to being a discontinuous stress measure, it can also be problematic in the context of automated driving since the driver is intended to be replaced by automation, leading to a decrease in driving behaviors (Lohani et al., 2019). Unlike subjective assessments and analysis of specific driving behaviors, physiological measures offer empirical evidence—objective and continuous—of the stress response (Plarre et al., 2011). Physiological measures thus offer a direct insight into the psychological and physiological adaptability of individuals dealing with stressful situations (Hancock and Warm, 1989). Finally, physiological measures remain relevant for monitoring driver stress during highly automated driving, during which drivers are not continuously in physical control of the vehicle.

Historically, stress responses have been compared to alarm states of the body, triggered by physical threats from the environment and intended to prepare the body for action (Selye, 1956). The alarm analogy provides a clear way of understanding the role of the physiological mechanisms that underlie stress responses and facilitate fast action-oriented reactions. Functionally, these mechanisms reflect a coactivation of autonomic components resulting in sympathetic autonomic stimulation and parasympathetic autonomic withdrawal, thus minimizing a vagal "braking" action on the motor system (Roelofs, 2017). Among physiological responses, cardiac measures are generally favored by researchers and practitioners for quantifying stress states. The most commonly used measures to explore cardiac activity are heart rate and Heart Rate Variability (HRV) (Alberdi et al., 2016). While heart rate focuses on contraction frequency, HRV is a measure of the time that elapses between contractions. The analysis of the time series of beat-to-beat intervals provides additional information since it reflects the heart's ability to adapt to changes by detecting and responding to stimuli over time (Acharya et al., 2006; Kim H. G. et al., 2018). The idea is that an individual with a low variability between heartbeats in a stressful context would have a low capacity to deal with stressful stimuli. In a driving context, a cardiac response to stressful stimuli is usually observed through an increase in heart rate (Healey and Picard, 2005; Lee et al., 2007; Cottrell and Barton, 2012; Guo et al., 2013; Zhao et al., 2014; Reimer et al., 2016; Rendon-Velez et al., 2016; Magana and Munoz-Organero, 2017; Antoun et al., 2018; Haouij et al., 2018; Khattak et al., 2018; Gotardi et al., 2019; Heikoop et al., 2019; Meesit et al., 2020) and a decrease in HRV (Lee et al., 2007; Yu et al., 2016; Heikoop et al., 2017; Magana and Munoz-Organero, 2017; Antoun et al., 2018; Rastgoo et al., 2019; Tavakoli et al., 2020; Zhao et al., 2020). Other physiological responses have also been studied as indexes of driver stress levels, such as changes in electrodermal activity (Healey and Picard, 2005; Cottrell and Barton, 2012; Pedrotti et al., 2014; Eisel et al., 2016; Morris et al., 2017; Ooi et al., 2018; Paredes et al., 2018; Zontone et al., 2020, 2021), breathing (Healey and Picard, 2005; Rendon-Velez et al., 2016; Balters et al., 2018; Haouij et al., 2018; Napoletano and Rossi, 2018; Heikoop et al., 2019; Zhao et al., 2020), blood pressure (Yamakoshi et al., 2008; Antoun et al., 2018), skin temperature (Yamakoshi et al., 2007, 2008; Zhao et al., 2020), muscle activation (Healey and Picard, 2005; Morris et al., 2017), pupil diameter (Pedrotti et al., 2014; Rendon-Velez et al., 2016; Zontone et al., 2021) and electrical brain activity (Kim S. et al., 2018; Halim and Rehan, 2020). Despite the numerous physiological responses studied, none of them has been validated as a definitive measure for identifying driver stress. Therefore, the use of a measure is often guided by practical and experimental design constraints (for a review of the advantages and disadvantages of physiological measures for assessing cognitive states in lab and real-world driving, see Lohani et al., 2019). Nevertheless, we believe that it is necessary for researchers and practitioners to base their measure selection decisions on both the practical constraints and the sensitivity to identify driver stress. Measure sensitivity refers to a measure's ability to discriminate between two levels of a psychological state (e.g., high and low stress) (Hughes et al., 2019). To date, the sensitivity of the driver stress measure has not been directly evaluated. Therefore, there is a need to specifically study the sensitivity of each physiological measure to driver stress to assist researchers and practitioners in measure selection.

Identifying Moderators of Driver Stress: A Theoretical Approach

Stress is a psycho-physiological state resulting from the influence of a stressor moderated by individual and ambient factors (Folkman and Lazarus, 1984; Matthews, 2002). In an automotive context, individual factors refer to the intrinsic characteristics of the driver (e.g., personality traits, demographic criteria), while ambient factors refer to the contextual effects (i.e., the circumstances in which a stressor operates).

Among the individual factors that may influence driver stress, age has probably been the most studied, particularly from a subjective perspective using self-report scales (Hartley and El Hassani, 1994; Simon and Corbett, 1996; Kloimüller et al., 2000). Despite these extensive investigations, the direction of the relationship between age and driver stress remains unclear. Indeed, some studies have found greater stress levels in older populations (Hill and Boyle, 2007) and explained this in terms of lower cognitive and physical abilities. Conversely, other studies have found lower stress levels in older populations (Langford and Glendon, 2002), which they have explained in part in terms of lower aggressiveness (Matthews et al., 1991; Westerman and Haigney, 2000) and more extensive driving experience (Gulian et al., 1990). Given the discrepancies at the subjective level, physiological measures provide objective ways of determining both the existence of the relationship and its direction. To our knowledge, only one study has found an effect of age on acute driver stress using physiological measures (Zhao et al., 2020). However, given the small number of participants included in this study (3 younger and 3 older), this effect deserves to be further explored. Like age, gender is an individual factor whose effect on driver stress is also debated. While some studies have found no effect of gender on driver stress using subjective scales (Wickens et al., 2015), others have reported higher stress levels in female drivers than male drivers based on cardiac (Guo et al., 2013) and hormone dosage measurements (Seeman et al., 1995).

In line with Hancock and Warm (1989), who recommended considering in stress studies both the demand imposed by the task and the type of environment, we suggest that automation (manual vs. autonomous) and stressor exposure duration (short vs. long-term) might be relevant factors when considering the driving task demand, while apparatus type (real vehicle vs. driving simulator) would make it possible to take account of the type of driving environment. We believe these three ambient factors to be of interest because they are either often debated in the literature (e.g., automation and apparatus), or have been the object of little direct study (e.g., stressor exposure duration).

Driving Automation

Interest in automated driving systems has grown over the last decade, in particular to compensate for the human errors in driving. More specifically in an automotive context, it is unclear whether a fully automated vehicle increases or reduces driver stress. Some authors have found positive effects of driving automation by reducing distress and enhancing driver attention (Funke et al., 2007), others have reported reduced driver stress coupled with a decrease in workload (Stanton and Young, 2005), while yet others have argued that autonomous driving increases driver stress due, in particular, to a lack of trust in the autonomous vehicle (Morris et al., 2017). Consequently, investigating this question would contribute to the development of automated driving systems adapted to the profiles of drivers and to given road situations.

Stressor Exposure Duration

The question regarding the existence of physiological differences between short and long periods of driving under acute stress has been little studied to date. A review of the literature came close to addressing this question by examining physiological responses to driver stress over short and long time periods (Antoun et al., 2017). However, due to the small number of studies collected, evidence of stress over a short time period was not revealed, thus reducing conclusions. The question therefore remains open.

Apparatus Type

With respect to the apparatus, the question of whether a driving simulator vs. a real vehicle is a valid way of studying internal driver states, such as stress, is unresolved. If the validity of simulators is confirmed, it is expected that observations made in a driving simulator will be equivalent to those made under real driving conditions. However, previous studies have reported contradictory results which make it difficult to draw clear conclusions. Taking the example of using mean heart rate to investigate validity, studies have shown a good level of correspondence between the simulator and the real road (Li et al., 2013). In contrast, other studies have found higher heart rates on real vehicles (Engström et al., 2005; Johnson et al., 2011). The fact that another study found both an absence of difference and a difference between the simulator and the real road depending on the driving situation, i.e., speed maintenance task and exposure to road hazards, respectively (Gemonet et al., 2021), further raises the question of the validity of the driving simulator for identifying driver stress in any driving situation.

Aims

We undertook a meta-analysis of the existing literature investigating driver stress, first to address, at a practical level, the difficulty researchers and practitioners have in selecting physiological measures for quantifying driver stress, and second, to gain insights into the relationship between driver stress and its moderators. The objectives were three-fold: (*i*) to investigate the sensitivity of each physiological measure used to quantify driver stress, (*ii*) to assess the moderating effect of the population type on driver stress, and (*iii*) to identify whether driver stress is influenced by ambient effects in the environment in which the driving task takes place.

METHODS

Search Strategy

This meta-analytical review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009).

Two investigators searched for articles in the electronic database, Google Scholar. The only limitation in terms of date was publication prior to February 2021. The following search terms were used: "{[(driver OR driving) AND (stress OR distress)] OR [(car) AND (stress OR distress)]}." These were then combined with additional terms related, first, to fields of research in which driver stress has been addressed: "psychological," "physiological," "behavior," "detection," "recognition" and, second, to the response of interest: "acute," "response," "change." In addition, a snowballing approach (Wohlin, 2014) was used to retrieve additional references. Duplicate records were systematically removed.

Each record was then screened (title, abstract and keywords) by the investigators in order to apply the eligibility criteria. The same procedure was carried out for the full-text articles. Any discrepancy between the investigators was resolved by discussion with a third investigator. The study selection process is described in **Figure 1** (PRISMA diagram).

Eligibility Criteria

We used the PICOS approach (Moher et al., 2015) to define the characteristics of studies eligible for inclusion in terms of population, interventions, comparators, outcomes and study design.

Population

Non-professional car drivers of all ages and genders, with no evidence of psychological or neurological disorders, were included.

Interventions

Stress interventions included driving tasks performed under high stress. Although the definition of "stress" or "high stress" is presumably a reflection of each author's particular standpoint, and the term has thus certainly been interpreted in many different ways, we decided to use Matthews' (2002) definition of driver stress to study similar stress interventions. Driver stress is thus interpreted as a psychological construct resulting from the stressful situation (involving stressors and ambient factors) and individual factors. Therefore, interventions in which driver stress was not a psychological construct but the product of physical action on the body were excluded. This was the case for stress interventions involving cold temperatures, pain, chronic illness, driving for long periods and monotonous driving periods.

Comparators

Comparators for the stress interventions were driving tasks performed under low stress.

Outcomes

All the included studies estimated driver stress based on physiological measures. All physiological outcomes were quantitatively reported as raw data or as means and standard deviations to allow the calculation of effect sizes. All physiological outcomes had been observed in at least three drivers.

Study Design

Only peer-reviewed quantitative physiological studies written in English were included in the analyses. All included studies contained a physiological measure also found in at least one other study to make it possible to compile the data required for a meta-analysis.

Data Extraction

For each included study, two investigators independently extracted the following data: demographic variables (sample size, mean age and gender ratio), ambient variables (apparatus, driving automation and stressor exposure duration), stress interventions and comparators (i.e., pairwise comparisons including a high stress intervention vs. a low stress intervention), statistical indices for the stress interventions and comparators (means and standard deviations) and type of physiological measure used.

When data was missing, the corresponding authors were contacted and asked for additional data. The WebPlotDigitizer software (Rohatgi, 2014) was also used to extract numerical values from the plot when numerical means and/or standard deviations were not reported.

For each included scientific paper in which driver stress was assessed in multiple population groups (e.g., older and younger participants), each pairwise comparison belonging to a given group was treated as a separate and independent study. As a result, and for the sake of clarity, we will use the term "study" in the following sections to refer to a pairwise comparison into a given group and not to the scientific paper from which it was derived. In addition, in studies that reported multiple stress interventions in the same population, the various stress interventions were averaged when raw data was available. This precaution was taken to avoid introducing an error due to the non-processed correlation between the condition effects estimated from multiple comparisons (Higgins et al., 2011). If raw data was not available, the highest-stress intervention was retained and the others were excluded. Although the strategy for selecting interventions is less recommended than combining interventions, it is generally difficult to obtain the raw data from each study, as would be required in order to compute the overall mean and standard deviation.

Meta-Analyses

Nine meta-analyses were conducted separately, one for each physiological measure. All analyses were carried out using JASP software (version 0.14.0.0). Due to different experimental designs and sample characteristics across included studies, we used random-effects models in an attempt to generalize our results beyond the studies included in our meta-analyses (Borenstein et al., 2010).

In keeping with previous studies that have tackled the issue of the sensitivity of physiological measures (Matthews et al., 2015; Hughes et al., 2019), we used effect size to determine the sensitivity of each measure of driver stress. Cohen's d effect size with 95% confidence intervals (95% CI) were first calculated for each study (i.e., for each pairwise comparison) based on the means, standard deviations and sample sizes (Cohen, 1988). Given the small sample sizes, Hedges' *g* was subsequently preferred to Cohen's *d* (Durlak, 2009). Hedges' *g* uses pooled weighted standard deviations instead of the pooled standard deviations used by Cohen's *d*. Mathematical equations

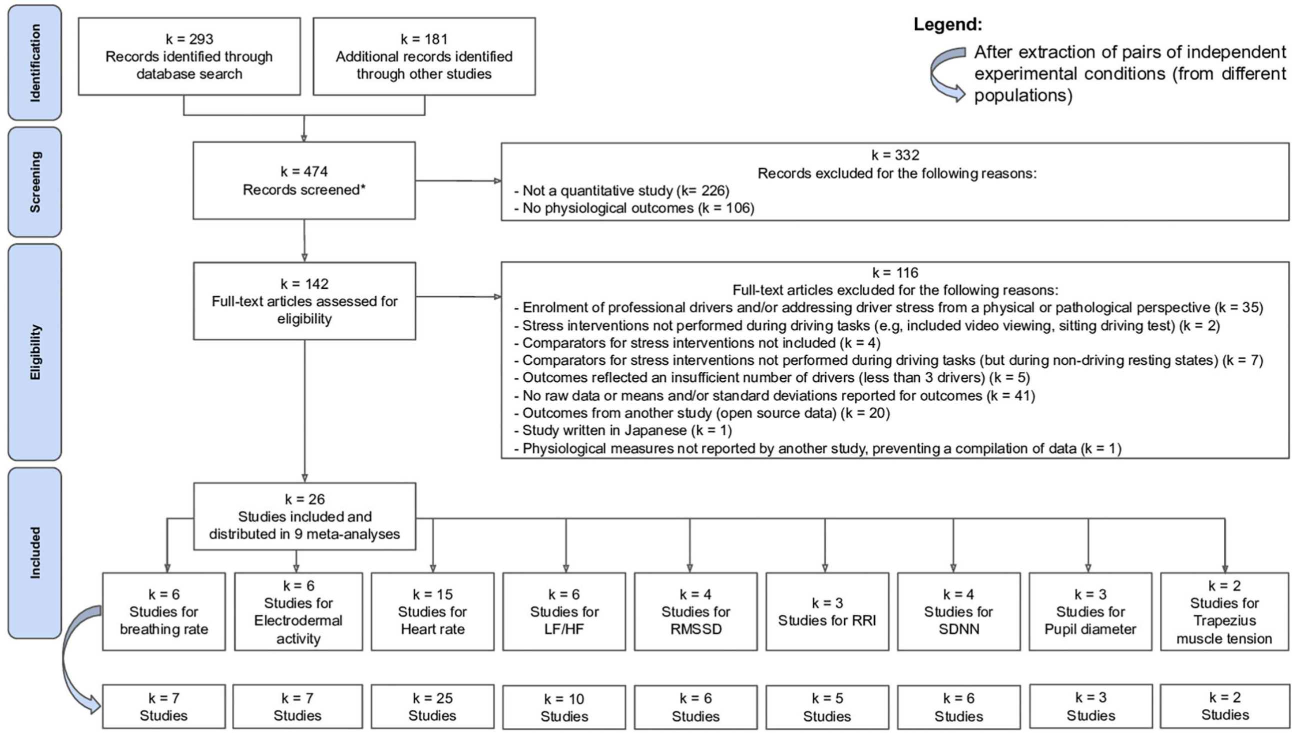


FIGURE 1 | PRISMA flowchart describing the methodology and search results. LF/HF, ratio of low frequency to high frequency; RMSSD, root mean square of successive differences among successive R-R normal intervals; RRI, means of R-R intervals; SDNN, standard deviation of normal R-R intervals.

used to compute effect size for each study are presented in the **Supplementary Material 1**. All effect sizes calculated for each study and corresponding to the same physiological measure were then aggregated to derive an overall summary effect size. A positive summary effect size indicated a positive effect of the stress intervention on all physiological measures except for HRV time-domain features (RRI, RMSSD and SDNN), for which a negative summary effect size suggested a positive effect of the stress intervention. Using Cohen's interpretation guidelines, the magnitude of the overall summary effect size was considered as small up to 0.2, medium up to 0.5, and large up to 0.8 (Cohen, 1988). The α level for significance was set at $p < 0.05$.

To quantify heterogeneity of the overall summary effect size, i.e., the inconsistency of effect sizes across a set of studies (Del Re, 2015), Q-statistic, I^2-statistic and τ^2 were explored. Q-statistic indicated the statistically significant presence of heterogeneity between effect sizes, I^2-statistic estimated the proportion of heterogeneity (low if $I^2 = 25\%$, moderate if $I^2 = 50\%$, large if $I^2 = 75\%$), and τ^2 referred to the absolute value of true variance across studies.

Publication bias was first assessed by visually inspecting the funnel plots. If an asymmetry was detected, a rank correlation test and an Egger's regression test (Egger et al., 1997) were run to assess the significance of the publication bias. Finally, the file drawer issue was assessed by Rosenthal's fail-safe N (Rosenthal, 1979). Fail-safe N refers to the number of studies that would have to be included in order to indicate that the stress intervention had no effect and that would be necessary for the meta-analysis to become non-significant. The file drawer problem was considered to be minor when the observed significance of fail-safe N was lower than the target significance level ($p = 0.05$), thus suggesting that the outcome of the meta-analysis was not affected by potential bias.

Moderator Analyses

Moderator analyses were undertaken if each measure met the three eligibility criteria: (1) significant summary effect size, (2) significant heterogeneity in summary effect size and (3) sufficient number of available studies ($k \geq 5$) to allow comparisons (Hughes et al., 2019). Meta-regressions were used when the factors studied were continuous variables, while subgroup analyses were conducted when the factors examined were categorical variables.

Individual Factors

Age and gender—two individual factors—were investigated by running meta-regressions to assess their moderating effect on driver stress.

Ambient Factors

The influence of three ambient factors on driver stress was studied by proceeding to subgroup analyses. These factors were: apparatus, driving automation and stressor exposure duration. In order to study the effect of apparatus type, a first subgroup was formed by pooling studies performed in a real vehicle while a second subgroup included studies performed in a driving simulator. Two independent analyses were then run to compute a summary effect for each subgroup. Finally, we analyzed whether the two summary effect sizes differed significantly, first by looking for overlaps between their confidence intervals and second by using a Wald-type test. The same procedure was repeated to explore driving automation and thus compare studies conducted in manual driving (first subgroup) and in fully autonomous driving (second subgroup). Again, the same procedure was used to investigate stressor exposure duration by comparing studies involving short-term exposure (first subgroup) and long-term exposure (second subgroup). The subgroups were formed by arbitrarily setting a threshold at 10 mins so that exposure times below the threshold comprised the first subgroup and exposure times above the threshold comprised the second subgroup.

RESULTS

Search Results

The primary search yielded 474 records. After screening each record, 332 abstracts were excluded in line with the eligibility criteria. The remaining 142 studies were then assessed for eligibility based on full-length articles. Finally, 26 references were included and distributed across 9 meta-analyses to permit independent exploration of 9 physiological measures (Healey and Picard, 2005; Schießl, 2008; Cottrell and Barton, 2012; Miller and Boyle, 2013; Manseer and Riener, 2014; Pedrotti et al., 2014; Zhao et al., 2014, 2020; Chen, 2015; Rendon-Velez et al., 2016; Yu et al., 2016; Heikoop et al., 2017, 2019; Magana and Munoz-Organero, 2017; Morris et al., 2017; Haouij et al., 2018; Khattak et al., 2018; Napoletano and Rossi, 2018; Ooi et al., 2018; Paredes et al., 2018; Gotardi et al., 2019; Rastgoo et al., 2019; Meesit et al., 2020; Tavakoli et al., 2020; Zontone et al., 2020, 2021) (**Figure 1**).

Characteristics of Studies

A qualitative review of the literature indicated that driver stress was indexed by breathing rate in 7 studies (156 drivers), electrodermal activity in 7 studies (187 drivers), heart rate in 25 studies (501 drivers), the ratio of Low-Frequency to High-Frequency heart rate variability (LF/HF) in 10 studies (140 drivers), the root mean square of successive differences among successive R-R normal intervals (RMSSD) in 6 studies (101 drivers), means of R-R intervals (RRI) in 5 studies (46 drivers), the standard deviation of normal R-R intervals (SDNN) in 6 studies (95 drivers), pupil diameter in 3 studies (83 drivers), and trapezius muscle tension in 2 studies (38 drivers). The characteristics of the studies included in the meta-analyses are detailed in **Supplementary Material 2**.

Meta-Analyses

The analyses indicated that several physiological measures changed significantly with stress interventions, thereby suggesting a change in drivers' stress state (**Table 1**). Indeed, heart rate [$g = 0.42$ (0.14 to 0.69), $p < 0.001$] and pupil diameter [$g = 0.46$ (0.02 to 0.90), $p < 0.05$] revealed significant moderate increases, while RRI, a time-domain feature of HRV, indicated a significant moderate decrease [$g = -0.42$ (-0.84 to 0.01), $p = 0.05$] when performing a high-stress driving task compared to a low-stress driving task. In contrast, no significant effects were observed between high-stress and low-stress driving for other measures, including breathing rate [$g = -0.27$ (-0.76 to 0.22), $p = 0.29$], electrodermal activity [$g = 0.96$ (-0.05 to 1.98), $p = 0.062$], LF/HF [$g = 0.60$ (-0.22 to 1.43), $p = 0.15$], RMSSD [$g = -0.06$ (-0.34 to 0.22), $p = 0.67$], SDNN [$g = -0.19$ (-0.47 to 0.10), $p = 0.20$] and trapezius muscle tension [$g = 0.04$ (-0.42 to 0.49), $p = 0.87$]. Among the measures that were found to be significantly sensitive to driver stress, i.e., heart rate, pupil diameter and RRI, none of them showed a real advantage over the others, as indicated by the overlap in their confidence intervals.

The Q-statistics indicated a significant heterogeneity between effect sizes for breathing rate [$Q = 20.3$, $p < 0.01$], electrodermal activity [$Q = 67.7$, $p < 0.001$], heart rate [$Q = 127.7$, p value < 0.001] and LF/HF ratio [$Q = 71.7$, p value < 0.001]. The degrees of heterogeneity for these measures, subsequently quantified using the I^2-statistic, were found to be moderate to large [Breathing rate: $I^2 = 66.0\%$ (16.1 to 94.8); Electrodermal activity: $I^2 = 94.0\%$ (84.4 to 98.9); Heart rate: $I^2 = 75.3\%$ (61.3 to 90.6); LF/HF: $I^2 = 89.6\%$ (77.0 to 96.9)]. Considering, first, the moderate to large degrees of uncertainty of I^2-statistics and, second, the amount of true variance between studies for these measures [Breathing rate: $\tau^2 = 0.23$; Electrodermal activity: $\tau^2 = 1.69$; Heart rate: $\tau^2 = 0.34$; LF/HF: $\tau^2 = 1.52$], we suspect that a large proportion of the observed variance reflected true heterogeneity.

Publication bias investigated by visually inspecting funnel plots for significant measures revealed no asymmetries (**Figure 2**). The absence of bias was then confirmed by standard rank correlation tests, Egger's regression tests, and fail-safe N analyses [Heart rate: Kendall's $\tau = 0.25$, $p = 0.10$, Egger: $z = 0.99$, $p = 0.32$, Fail-safe N = 350, $p < 0.001$; RRI: Kendall's $\tau = -0.32$, $p = 0.45$, Egger: $z = -1.26$, $p = 0.21$, Fail-safe N = 4, $p < 0.05$; Pupil diameter: Kendall's $\tau = 0.33$, $p = 1.00$, Egger: $z = 1.05$, $p = 0.29$, Fail-safe N = 5, $p < 0.01$].

Moderator Analyses

To determine the extent to which physiological measures are sensitive to individual and ambient factors, we carried out a series of moderator analyses using subgroups and meta-regressions. Only heart rate met the three eligibility criteria required to conduct moderator analyses: significant summary effect size [$g = 0.42$ (0.14 to 0.69), $p < 0.001$], significant heterogeneity in summary effect size ($Q = 127.7$, $p < 0.001$), and sufficient number of available studies ($k = 25 \geq 5$).

Individual Factors

The moderating effects of age and gender on driver stress were explored (**Table 2**). Meta-regressions revealed no effect of age ($\beta = -0.015$, $p = 0.22$) or gender ($\beta = -0.003$, $p = 0.48$).

Ambient Factors

We assessed the moderating effects on driver stress of three ambient factors: apparatus, driving automation and stressor

TABLE 1 | Outcomes of the meta-analyses.

Physiological measure	Sample size		Heterogeneity			Global effect size		
	k	N	*Q*-statistic	*I*²-statistc (%)	*t*²	Hedges' g	95%CI	*p*-value
Breathing rate	7	156	**20.3****	66.0	0.23	−0.27	[−0.76; 0.22]	0.29
Electrodermal activity	7	187	**67.7*****	94.0	1.69	0.96	[−0.05, 1.98]	0.062
Heart rate	25	501	**127.7*****	75.3	0.34	0.42	[0.14; 0.69]	**<0.001*****
LF/HF	10	140	**71.7*****	89.6	1.52	0.60	[−0.22; 1.43]	0.15
RMSSD	6	101	0.60	0.00	0.00	−0.06	[−0.34; 0.22]	0.67
RRI	5	46	6.51	0.00	0.00	−0.42	[−0.84; 0.01]	**0.05***
SDNN	6	95	0.99	0.00	0.00	−0.19	[−0.47; 0.10]	0.20
Pupil diameter	3	83	1.85	20.4	0.038	0.46	[0.02; 0.90]	**<0.05***
Trapezius muscle tension	2	38	0.11	0.00	0.00	0.04	[−0.42; 0.49]	0.87

*LF/HF, ratio of low frequency to high frequency; RMSSD, root mean square of successive differences among successive R-R normal intervals; RRI, means of R-R intervals; SDNN, standard deviation of normal R-R intervals; k, number of studies; N, number of drivers; Q, I² and τ², statistics used to evaluate heterogeneity of variance; Hedges' g, statistic used to calculate effect size for small sample size; CI, confidence interval; p-value, level of significance. *p< 0.05, **p< 0.01, ***p< 0.001.*

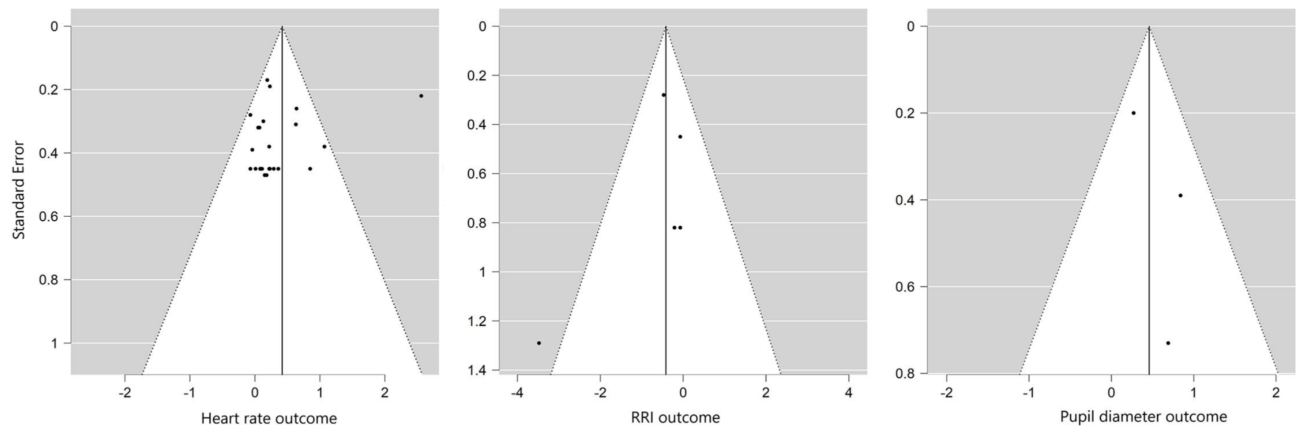


FIGURE 2 | Visualization of funnel plots.

exposure duration (**Table 3**). The first ambient factor tested was the apparatus. No significant change in heart rate was observed between driving tasks performed in the real-vehicle and driving tasks performed in a driving simulator [$g_{\text{Real}} = 0.37$ (0.00 to 0.74), $g_{\text{Simulator}} = 0.41$ (0.11 to 0.71)], as revealed by the overlapping of their confidence intervals. These observations were reinforced by the Wald-type test, which did not indicate any significant difference between the two summary effect sizes ($z_{\text{Apparatus}} = 0.44$, $p = 0.66$).

The second ambient factor we assessed was driving automation. Although heart rate showed a greater overall effect size when stress intervention was performed in manual driving [$g_{\text{Manual}} = 0.47$ (0.16 to 0.77)] compared to fully autonomous driving [$g_{\text{Fullyautonomous}} = 0.09$ (−0.33 to 0.51)], the overlap in the confidence intervals suggested that the difference was not statistically significant. In addition, the results of the Wald-type test indicated similar summary effect sizes between manual and autonomous driving ($z_{\text{Automation}} = 0.87$, $p = 0.38$).

The third ambient factor assessed was the stressor exposure duration. No significant cardiac difference was noticed between short and long-term stress exposure [$g_{\text{short}} = 0.44$ (0.10 to 0.79), $g_{\text{Long}} = 0.22$ (−0.05 to 0.49)]. The lack of significance was indeed supported by the Wald test result ($z_{\text{Duration}} = 0.31$, $p = 0.76$).

DISCUSSION

To our knowledge, these are the first meta-analyses to investigate (*i*) the sensitivity of each physiological measure in quantifying driver stress, and the moderating effect of (*ii*) population type and (*iii*) driving ambient on driver stress. The main finding is that moderate physiological changes were initiated by stress interventions, suggesting that heart rate, RRI—a time-domain HRV feature—and pupil diameter are sensitive measures for quantifying driver stress. Driver stress indexed by heart rate showed no moderating effect of age, gender, apparatus, driving automation or stressor exposure duration. Below, we provide a summary and interpretations of the results, discuss implications for future research and present the main limitations of the reported work.

TABLE 2 | Outcomes of individual factors.

Individual factor	physiological measure	Sample Size		Heterogeneity			Global effect size			
		k	*N*	*Q*-statistic	I^2-statistc (%)	t^2	Hedges' g	β	SE	*p*-value
Age										
	Heart Rate	24	491	**114.2****	75.1	0.34	0.91	–0.015	0.45	0.22
Women										
	Heart Rate	24	491	**124.4****	76.9	0.36	0.55	–0.003	0.24	0.48

*k, number of studies; N, number of drivers; I^2, τ^2 and Q, statistics used to evaluate heterogeneity of variance. Hedges' g, statistic used to calculate effect size for small sample size; β, non-standardized beta coefficient; SE, standard error; p-value, level of significance. **p< 0.01.*

TABLE 3 | Outcomes of ambient factors.

Ambient factor	Physiological measure	Sample size		Heterogeneity			Global effect size			
		k	N	*Q*-statistic	I^2-statistc (%)	t^2	Hedges' g	95%CI	SE	*p*-value
Apparatus										
Real vehicle	Heart Rate	16	256	**97.6****	73.4	0.41	0.37	[0.00, 0.74]	0.19	0.053
Driving simulator	Heart Rate	9	245	**24.4****	57.4	0.11	0.41	[0.11, 0.71]	0.15	**<0.01****
Automation										
Manual	Heart Rate	22	457	**124.0****	78.0	0.39	0.47	[0.16, 0.77]	0.16	**<0.01****
Fully autonomous	Heart Rate	3	44	0.2	0.0	0.00	0.09	[–0.33, 0.51]	0.21	0.67
Duration										
Short-term	Heart Rate	17	391	**108.3****	79.9	0.40	0.44	[0.10, 0.79]	0.18	**<0.05***
Long-term	Heart Rate	8	110	**15.1***	0.0	7.57e-6	0.22	[–0.05, 0.49]	0.14	0.115

*k, number of studies; N, number of drivers; I^2, τ^2 and Q, statistics used to evaluate heterogeneity of variance; Hedges' g, statistic used to calculate effect size for small sample size; CI, confidence interval; SE, standard error; Apparatus, real vehicle vs. driving simulator; Automation, manual vs. autonomous driving; Duration, short-term vs. long-term exposure to the stressor; p-value, level of significance. *p< 0.05, **p < 0.01.*

Summary and Interpretation of the Results

Considering the overall effect sizes and their confidence intervals in order to judge the significance of an effect, and thus the sensitivity of a measure, we identified three physiological measures that are sensitive enough to quantify driver stress, namely heart rate, RRI and pupil diameter. The fact that both heart rate and RRI are both sensitive is consistent since heart rate is derived from RRI. It should be noted that of the three sensitive physiological measures (i.e., heart rate, RRI and pupil diameter), none was found to have a significant advantage over any other in identifying driver stress. While these three measures showed sensitivity to driver stress, the other measures did not (i.e., breathing rate, electrodermal activity, LF/HF, RMSSD, SDNN and trapezius muscle tension). However, this does not mean that they are not sensitive. At this stage, we cannot conclude about the lack of sensitivity of these measures. It is indeed possible that the sample size for each of these measures is too small and/or presents too much heterogeneity across studies, which would prevent revealing a sensitivity to driver stress.

Only heart rate warranted moderator analysis because it was the only measure that met all the eligibility criteria. However, individual moderators (age, gender) and ambient moderators (apparatus, driving automation, stressor exposure duration) did not reveal any significant change in heart rate. Despite this, it is very likely that there are moderators of the stress response given the considerable heterogeneity (i.e., high values of Q, I^2 and τ^2) observed in the effect sizes. Possible explanations regarding the lack of physiological change are provided below.

Individual Modulators

Age and Gender

Although it is well established that individual factors have an impact on stress appraisal (Matthews, 2002), the results regarding the direction of the relationship between individual factors and driver stress have often been contradictory. For example, studies have shown greater stress levels in older populations (Hill and Boyle, 2007), while others have observed lower stress levels in older populations (Langford and Glendon, 2002). Therefore, the aggregation of studies with opposite results in the same meta-analysis could explain our findings about the lack of an age effect on driver stress. Nonetheless, this does not mean that there is no real moderating effect of age. Indeed, the driving experience, closely linked to age (Gulian et al., 1990), can influence the driver stress response, as observed through the stronger correlations between age and all dimensions of driver stress (DBI scales) when driving experience is statistically controlled (Westerman and

Haigney, 2000). Also, cognitive decline has been mentioned as a possible explanation for greater stress levels in older populations, which is highlighted, in particular, by a drop in "alertness and anticipation" and an increase in "driving dislike" with age (Westerman and Haigney, 2000). Therefore, the unifactorial approach (i.e., investigating factors one by one) might mask the true effect of moderating factors (e.g., age and gender, lack of experience or negative experiences, awareness of cognitive decline) by not taking account of their interdependence. This is in line with Matthews' (2002) transactional theory of driver stress, according to which driver stress is the result of transactional relationships between several factors.

Ambient Modulators

Apparatus

Although stress studies conducted in a driving simulator offer a more controlled and safe approach, they might nevertheless be poorly representative of the stress experienced under real and ecological conditions. Our results *a priori* seem to contradict this criticism since they suggest that stress induced in a driving simulator and measured by heart rate is indeed representative of stress experienced in real conditions. Indeed, the lack of change in heart rate between driving simulator studies and real vehicle driving studies was observed through similar overall effect sizes, similar standard errors and a non-significant Wald-type test. However, the significant heterogeneity in effect sizes, observed in both simulator and real-road studies, indicates that additional factors explain the overall effect size. We believe that these factors are related to differences in experimental designs, and in particular in the stressful stimuli used. In addition, it cannot be excluded that the nature of the stimuli used and the experimental designs also differ between studies conducted on driving simulators and in real-vehicles. Thus, we can legitimately ask whether the internal driver states we measure in driving simulators and in real road conditions are the same, and if the response to stressful stimuli in real car driving is not shaped by additional safety concerns, among other factors. This is why Milleville-Pennel and Charron (2015) raised the question: "*Can we consider that the same cognitive functions are involved in simulated driving and in real car driving?.*" Furthermore, previous studies have compared internal driver states (not exclusively stress) in simulated and real-world driving using the same stimuli and have measured these states using heart rate (Engström et al., 2005; Johnson et al., 2011; Li et al., 2013; Gemonet et al., 2021). However, no consensus has been reached due to conflicting results. Given both our results and the discrepancy between results in the literature, we recommend further investigating driver stress in both simulated and real vehicle driving using experimental designs that are as similar as possible, i.e., including the same hazardous or stressful stimuli, same driving environment and same participants when doing driving simulator validation studies.

Automation

The lack of difference in measures of heart rate between manual and autonomous driving—indicated by a non-significant Wald-type test—indicates *a priori* that driver stress is not influenced by driving automation. Nonetheless, the effect size of stress interventions was significant in manual driving ($g = 0.47$, $p < 0.01^{**}$), while it was non-significant in autonomous driving ($g = 0.09$, $p = 0.67$). Taken together, the lack of difference observed between manual and autonomous driving may be due to the small number of included studies that investigated autonomous driving (k = 3). Although no reliable conclusion concerning the possible influence of driving automation on driver stress can be provided at this stage, further investigations of driver stress in autonomous driving are strongly recommended to confirm or refute this lack of effect. In cases where additional studies confirm this lack of effect, it would be interesting to explore the sources. Below, we put forward potential explanations for the lack of an effect of autonomous driving that can be considered as avenues of investigation. First, such a lack of effect may be due to the different nature of the stressors, i.e., more arousing and demanding in terms of cognitive and motor skills for manual driving than for automated driving. Second, it may also be explained by a reduction in driver stress during autonomous driving. This explanation would be consistent with the hypothesis of reduced vulnerability to stress during autonomous driving and related to the decrease in workload (Stanton and Young, 1998, 2005). Third, the lack of effect of stress interventions may also be due to drivers' level of experience with automated driving systems and their trust. As evidence of this, a relationship has previously been found between reported trust in autonomous driving and physiological stress (Morris et al., 2017). Fourth, heart rate may not be a suitable indicator for detecting stress in autonomous driving. Therefore, it would be interesting to consider alternative measures, such as LF/HF ratio (Heikoop et al., 2017) and electrodermal activity (Zontone et al., 2020), both of which have already been used for stress detection purposes during autonomous driving.

Duration

The lack of change in heart rate between short-term and long-term driving—highlighted by a non-significant Wald-type test—suggests that the sensitivity of heart rate is not modulated by the stressor exposure duration. However, the effect size of stress interventions was significant in short-term driving ($g = 0.44$, $p < 0.05^{*}$), whereas it was non-significant in long-term driving ($g = 0.22$, $p = 0.115$). Although additional studies would be necessary to draw definitive conclusions concerning the existence of cardiac differences depending on the duration of driving under stress conditions, the disparity of the results nevertheless enables us to put forward a first hypothesis. Indeed, it is likely that our findings reflect the effect of the nature of the stressors manipulated within each subgroup (short-term and-long-term) and not the effect of the stressor exposure duration and therefore the measurement time. We believe that event-related and intense stressors are more likely to be studied over short time periods than more diffuse and moderate stressors, which would require longer measures in order to be detected by cardiac sensors. Consequently, in the future, it would be interesting to study the same stressors (i.e., same nature and intensity) while varying only the cardiac measurement time. This would also address the question raised by Antoun et al. (2017) about the existence of a threshold effect beyond which driving in a given context would become significantly more stressful. For exploratory purposes, a driving time cut-off of 10 mins was arbitrarily set when forming the subgroups and it is possible that other values might be more appropriate for highlighting a potential moderating effect of stressor exposure duration on driver stress.

Implications for Future Research and Practice

Our results aim to shed light on driver stress-sensitive measures in order to assist researchers and practitioners in their measurement decisions. Based on our findings, three physiological measures were found to be sensitive to driver stress, namely heart rate, RRI and pupil diameter. Nonetheless, we recommend that readers interpret our results (i.e., the magnitude of the effects) in the context in which driver stress was manipulated in the included studies. Indeed, as Mehler et al. (2012) suggested, the sensitivity of measures may vary depending on the specific tasks and individual states considered. In addition, we encourage further investigation of the other measures used, which may not have been able to reveal their potential sensitivity in our study, in part because of the limited number of studies and/or failure of studies to meet eligibility criteria.

Considerations for future research and practice arise mainly from the results of sensitivity and moderator analyses. We found, first, that some measures did not exhibit sensitivity to stress and that the studied factors did not highlight a moderating effect on stress despite the large heterogeneity in effect sizes. As a result, we recommend that researchers and practitioners interested in exploring driver stress adopt a 3-step approach in order to optimize the observation of both physiological change reflecting sensitivity and of moderating effects, and, more generally, to improve the understanding of driver stress. The 3-step approach consists of: (1) using multiple measures, (2) combining measurement modalities (triangulation approach), and (3) analyzing how factors (stressors and moderators) interact (multifactorial approach). Below, we advocate these principles for driver stress investigations, although they can also be applied to the exploration of other psycho-physiological and cognitive states.

Using Multiple Measures

First, researchers and practitioners should use multiple measures to ensure that the physiological changes induced by stressors are also actually observed. This approach would compensate for the failure of some measures in some individuals or in some study contexts. For example, Healey and Picard (2005) pointed out that the electrodermal response may differ among drivers due to variations in the number of sweat glands on the palms. The question of the reliability of pupil diameter to index driver stress also arises in real road contexts, where the measure can be disturbed by many uncontrollable factors, such as light variation and driver's verbal output (Recarte and Nunes, 2003). According to Mehler et al. (2012), no single physiological measure would provide optimal sensitivity for capturing a given state in all types of tasks. Second, using multiple measures in combination would permit a more reliable identification of driver stress. Indeed, Bernardi et al. (2000) supported the analysis of combined measures after observing the influence of breathing on HRV during simple mental and verbal activities. More specifically in an automotive context, the influence of driver stress resulting from a combination of physiological measures has also been investigated (Ollander et al., 2016). The authors found that combining cardiac, electrodermal and respiratory signals made it possible to distinguish between resting and driving, while combining cardiac and respiratory signals helped distinguish between low-stress driving and high-stress driving (Ollander et al., 2016). Third, the use of multiple measures and features would also provide information about the sympathovagal balance, thus improving knowledge of the psychophysiological mechanisms underlying stress states. Some measures and features reflect the activity of both autonomic components, while others mainly reflect the activity of one of the two components. This knowledge is also particularly interesting for remediation strategies, given that Respiratory Sinus Arrhythmia (RSA) mainly reflects the parasympathetic component (Berntson et al., 1993), that a low RSA and anxiety are related (Thayer et al., 1996) and that it has proved possible to progressively increase RSA using breathing and biofeedback techniques (Climov et al., 2014).

Triangulation Approach

In the same way as other works which have previously reviewed studies of stress (Alberdi et al., 2016), and driver stress in particular (Rastgoo et al., 2018; Chung et al., 2019), we advocate the joint use of physiological, subjective, and behavioral measures to explore stress in driving. This approach, also called triangulation (Denzin, 1978), permits the accurate observation of a common phenomenon and enriches its explanation (Jick, 1979). Since such an approach captures the multidimensional responses to stress (Matthews, 2002) at the physiological, behavioral, emotional and cognitive levels, it will help us differentiate between the various stress states experienced by drivers. This will then make it possible to derive stress-sensitive driver profiles (Pesle et al., 2018) and design driver stress detection systems (Rastgoo et al., 2018).

Multifactorial Approach

Our results showed no modulating effect of the studied factors (age, gender, apparatus, driving automation, and stressor exposure duration). As suggested above, these findings may be partly due to our univariate approach, which considered each factor independently. This statement is supported by a recent study in which an effect of age on driver stress was found using a multivariate approach (i.e., Principal Component Analysis of physiological measures) (Zhao et al., 2020). This type of approach has been supported by a number of different studies which have observed dependencies between driver stress and various individual and ambient factors, such as personality, mood, coping strategies, age, gender, driving experience, time of day in relation to the circadian rhythms (Langford and

Glendon, 2002; Pesle et al., 2018). Our findings, alongside those of previous studies, support the idea that the multivariate approach advocated by Matthews et al. (2017) if we are to achieve a holistic understanding of the moderators (individual and ambient), stressors and outcomes of driving. Nonetheless, this type of approach remains difficult to implement. In this context, the multivariate approach should systematically call on theoretical support, such as the T^2SO (Time-Trait-Stressors-Outcome) framework proposed by Matthews et al. (2017), to facilitate understanding and test multivariate theories of driver stress. In addition, the use of computational techniques would facilitate the implementation of a multifactorial approach.

LIMITATIONS

Several Limitations Should Be Acknowledged

Small Number of Studies

Although the random-effects models used for our meta-analyses were designed to permit us to generalize our results beyond the included studies (Borenstein et al., 2010), the small number of studies nevertheless limits the scope of our interpretations. Given the small number of studies, moderator analyses could be performed for only one stress-sensitive physiological measure; namely, heart rate. Therefore, it cannot be excluded that the results and interpretations of the moderator analyses are dependent on the physiological measure used, in this case heart rate. Interpretations of each moderator are also limited by the small number of studies within some moderator subgroups. This reflects the fact that driver stress has not been sufficiently investigated under specific driving conditions (e.g., autonomous driving). One reason for the small number of studies included in meta-analyses is the exclusion of driver stress studies that used various algorithms to combine physiological signals (Singh et al., 2011; Lanatà et al., 2014; Dobbins and Fairclough, 2018; Bitkina et al., 2019; Hadi et al., 2019). Indeed, we focused on a univariate approach to examine the sensitivity of independent physiological measures. Another major reason is the lack of information about the stress interventions in the studies (e.g., mean and/or standard deviation).

Use of Different Stressors

As driver stress has been interpreted in different ways by authors, many stress interventions have been collected across studies (e.g., heavy traffic, complex driving maneuvers, surprising events). Therefore, the effect sizes could be identified more precisely if comparison groups included only highly similar stressors. The wide variety of experimental designs found in the studies did not allow us to achieve such granularity.

Highlight Sensitivity of Physiological Measures to Driver Stress, but Not Selectivity (or Specificity)

The current study demonstrated the sensitivity—and not the selectivity—of various physiological measures to driver stress. Sensitivity refers to the capacity of an instrument to detect changes in a given task or situation, whereas selectivity refers to the sensitivity of an instrument only to differences in one state (e.g., stress state) and not changes in other states (e.g., mental workload) (O'Donnell and Eggemeier, 1986; Matthews et al., 2015). It is therefore entirely possible that the physiological measures found to be sensitive to driver stress in this study are also sensitive to other psycho-physiological and cognitive states of the driver. Several factors (i.e., not only stressors) would thus influence the autonomic nervous system responses. Such observations would suggest a lack of selectivity of the physiological measures to driver stress when the measures are used alone and independently, i.e., without combining measures. In favor of this assumption, let's take the example of driver stress-sensitive heart rate. Zontone et al. (2020) noted a systematic difference in heart rate between manual and autonomous driving under all conditions (stress and control), leading them to believe that additional factors, unrelated to stress, were responsible for the changes in heart rate. One of the most likely explanations for these changes in heart rate is the significant influence of motor activity during manual driving. Another possible explanation is that mental workload influences cardiac response, which would consequently be reduced with automation (Stanton and Young, 1998; Young and Stanton, 2002). In addition, Parent et al. (2019) suggested that stress and mental workload would have similar sources and effects. Given these common characteristics, the use of a single physiological measure, in this case heart rate, might be limited in its ability to infer a specific state (e.g., stress state) when several factors interplay (e.g., stress, mental workload, motor activity). The current study investigated the physiological measures alone and independently, therefore it meets the criterion of sensitivity of the physiological measures to driver stress but not selectivity. We believe that the investigation of the selectivity of physiological measures to driver stress can only be done by considering multiple driver states, including multiple measures, combining multiple measurement modalities, and performing an analysis of multiple explanatory factors. Although this approach is highly challenging to implement, we have good reason to believe that the multivariate approach is the key to distinguishing each driver state, including driver stress. In this sense, previous research has shown the specificity of autonomic nervous system responses to basic emotions when these emotions were examined using multivariate analyses (Stephens et al., 2010). Given the importance of emotions (e.g., anger, fear) in the driver's stress response, multivariate analyses might be a powerful tool to enable isolating stress from other psychophysiological and cognitive states. Computational techniques (e.g., preprocessing, feature selection, machine learning) and neuroimaging techniques, which have recently been shown to differentiate stress from workload (Parent et al., 2019), might also contribute to distinguishing all these states.

CONCLUSION

This research relied on an empirical approach that aggregates results from the literature to quantify the sensitivity of physiological measures to driver stress. The results showed that heart rate, RRI and pupil diameter were sensitive enough to permit this. We believe that these findings could provide initial support for researchers and practitioners when deciding which physiological measures to use to quantify stress while driving.

Future studies involving these measures, as well as HRV features, electrodermal activity, breathing rate and trapezius muscle tension, are necessary to draw conclusions about their (lack of) sensitivity for quantifying driver stress. Given the growing interest in achieving early detection, we recommend using multiple physiological measures in order to ensure and enhance the observation of stressor-induced physiological changes. Indeed, the design of corrective or assistance solutions that specifically target driver stress and that would be activated as soon as stress emerges would be of interest in terms of safety and comfort. In addition, in order to promote a broad understanding of driver stress involving stressors, modulators and outcomes, we recommend a triangulation-based methodology (using subjective, behavioral and physiological measures) combined with a multifactorial approach (studying several factors simultaneously and jointly). Finally, functional neuroimaging studies should be performed to explore the neurophysiological correlates underlying driver stress states and thus provide additional insights into these states.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

LK, SD, and JN contributed to the meta-analysis process, contributed to manuscript revision, read, approved the submitted version, and collected the data. LK organized the database, wrote the first draft of the manuscript, and performed the statistical analysis. JN revised the statistical analysis. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnrgo.2021.756473/full#supplementary-material

REFERENCES

Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M., and Suri, J. S. (2006). Heart rate variability: a review. *Med. Biol. Eng. Comput.* 44, 1031–1051. doi: 10.1007/s11517-006-0119-0

Alberdi, A., Aztiria, A., and Basarab, A. (2016). Towards an automatic early stress recognition system for office environments based on multimodal measurements: a review. *J. Biomed. Inform.* 59, 49–75. doi: 10.1016/j.jbi.2015.11.007

Antoun, M., Ding, D., Bohn-Goldbaum, E. E., Michael, S., and Edwards, K. M. (2018). Driving in an urban environment, the stress response and effects of exercise. *Ergonomics* 61, 1273–1281. doi: 10.1080/00140139.2018.1468494

Antoun, M., Edwards, K. M., Sweeting, J., and Ding, D. (2017). The acute physiological stress response to driving : a systematic review. *PLoS ONE* 12:e0185517. doi: 10.1371/journal.pone.0185517

Balters, S., Murnane, E. L., Landay, J. A., and Paredes, P. E. (2018). "Breath booster! exploring in-car, fast-paced breathing interventions to enhance driver arousal state," in *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare* (pp. 128–137).

Bernardi, L., Wdowczyk-Szulc, J., Valenti, C., Castoldi, S., Passino, C., Spadacini, G., et al. (2000). Effects of controlled breathing, mental activity and mental stress with or without verbalization on heart rate variability. *J. Am. Coll. Cardiol.* 35, 1462–1469. doi: 10.1016/S0735-1097(00)00595-7

Berntson, G. G., Cacioppo, J. T., and Quigley, K. S. (1993). Respiratory sinus arrhythmia: autonomic origins, physiological mechanisms, and psychophysiological implications. *Psychophysiology* 30, 183–196. doi: 10.1111/j.1469-8986.1993.tb01731.x

Bitkina, O. V., Kim, J., Park, J., Park, J., and Kim, H. K. (2019). Identifying traffic context using driving stress: A longitudinal preliminary case study. *Sensors.* 19:2152. doi: 10.3390/s19092152

Borenstein, M., Hedges, L. V., Higgins, J. P., and Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res. Synth. Methods* 1, 97–111. doi: 10.1002/jrsm.12

Chen, Y. (2015). Detecting drivers' stress in emergency using physiological sensors. *Appl. Mechan. Mater.* 742, 150–157. doi: 10.4028/www.scientific.net/AMM.742.150

Chung, W. Y., Chong, T. W., and Lee, B. G. (2019). Methods to detect and reduce driver stress: a review. *Int. J. Automot. Technol.* 20, 1051–1063. doi: 10.1007/s12239-019-0099-3

Climov, D., Lysy, C., Berteau, S., Dutrannois, J., Dereppe, H., Brohet, C., et al. (2014). Biofeedback on heart rate variability in cardiac rehabilitation: practical feasibility and psycho-physiological effects. *Acta Cardiol.* 69, 299–307. doi: 10.1080/AC.69.3.3027833

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. London: Routledge.

Cottrell, N. D., and Barton, B. K. (2012). The impact of artificial vehicle sounds for pedestrians on driver stress. *Ergonomics* 55, 1476–1486. doi: 10.1080/00140139.2012.724452

Del Re, A. C. (2015). A practical tutorial on conducting meta-analysis in R. *Quant. Meth. Psychol.* 11, 37–50. doi: 10.20982/tqmp.11.1.p037

Denzin, N. K. (1978). *The Research Act: A Theoretical Introduction to Sociological Methods*. New York, NY: McGraw-Hill.

Dobbins, C., and Fairclough, S. (2018). Signal processing of multimodal mobile lifelogging data towards detecting stress in real-world driving. *IEEE Trans. Mobile Comput.* 18, 632–644. doi: 10.1109/TMC.2018.2840153

Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *J. Pediatr. Psychol.* 34, 917–928. doi: 10.1093/jpepsy/jsp004

Egger, M., Smith, G. D., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315, 629–634. doi: 10.1136/bmj.315.7109.629

Eisel, M., Nastjuk, I., and Kolbe, L. M. (2016). Understanding the influence of in-vehicle information systems on range stress–Insights from an electric vehicle field experiment. *Transport. Res. part F: Traffic Psychol. Behav.* 43, 199–211. doi: 10.1016/j.trf.2016.10.015

Emo, A. K., Matthews, G., and Funke, G. J. (2016). The slow and the furious: anger, stress and risky passing in simulated traffic congestion. *Transport. Res. part F: Traffic Psychol. Behav.* 42, 1–14. doi: 10.1016/j.trf.2016.05.002

Engström, J., Johansson, E., and Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transport. Res. part F: Traffic Psychol. Behav.* 8, 97–120. doi: 10.1016/j.trf.2005.04.012

Espejo, E. P., Ferriter, C. T., Hazel, N. A., Keenan-Miller, D., Hoffman, L. R., and Hammen, C. (2011). Predictors of subjective ratings of stressor severity: the effects of current mood and neuroticism. *Stress Health* 27, 23–33. doi: 10.1002/smi.1315

Folkman, S., and Lazarus, R. S. (1984). *Stress, Appraisal, and Coping*. New York, NY: Springer Publishing Company.

Funke, G., Matthews, G., Warm, J. S., and Emo, A. K. (2007). Vehicle automation: a remedy for driver stress?. *Ergonomics* 50, 1302–1323. doi: 10.1080/00140130701318830

Ge, Y., Qu, W., Jiang, C., Du, F., Sun, X., and Zhang, K. (2014). The effect of stress and personality on dangerous driving behavior among Chinese drivers. *Accid. Anal. Prevent.* 73, 34–40. doi: 10.1016/j.aap.2014.07.024

Gemonet, E., Bougard, C., Masfrand, S., Honnet, V., and Mestre, D. R. (2021). Car drivers coping with hazardous events in real vs. simulated situations: declarative, behavioral and physiological data used to assess drivers' feeling of presence. *PLoS ONE* 16:e0247373. doi: 10.1371/journal.pone.0247373

Glendon, A. I., Dorn, L., Matthews, G., Gulian, E., Davies, D. R., and Debney, L. M. (1993). Reliability of the driving behaviour inventory. *Ergonomics* 36, 719–726. doi: 10.1080/00140139308967932

Gotardi, G. C., Polastri, P. F., Schor, P., Oudejans, R. R. D., Van Der Kamp, J., Savelsbergh, G. J. P., et al. (2019). Adverse effects of anxiety on attentional control differ as a function of experience: a simulated driving study. *Appl. Ergon.* 74, 41–47. doi: 10.1016/j.apergo.2018.08.009

Gulian, E., Glendon, A. I., Matthews, G., Davies, D. R., and Debney, L. M. (1990). The stress of driving: a diary study. *Work Stress* 4, 7–16. doi: 10.1080/02678379008256960

Gulian, E., Matthews, G., Glendon, A. I., Davies, D. R., and Debney, L. M. (1989). Dimensions of driver stress. *Ergonomics* 32, 585–602. doi: 10.1080/00140138908966134

Guo, W., Brennan, D., and Blythe, P. T. (2013). "Detecting older drivers' stress level during real-world driving tasks," in *International Conference on Applied Psychology and Behavioral Sciences, Paris*, p. 307.

Hadi, W. E., El-Khalili, N., AlNashashibi, M., Issa, G., and AlBanna, A. A. (2019). Application of data mining algorithms for improving stress prediction of automobile drivers: a case study in Jordan. *Comput. Biol. Med.* 114:103474. doi: 10.1016/j.compbiomed.2019.103474

Halim, Z., and Rehan, M. (2020). On identification of driving-induced stress using electroencephalogram signals: a framework based on wearable safety-critical scheme and machine learning. *Informat. Fus.* 53, 66–79. doi: 10.1016/j.inffus.2019.06.006

Hancock, P. A., and Desmond, P. A. (2001). *Stress, Workload, and Fatigue*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Hancock, P. A., and Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Hum. Fact.* 31, 519–537. doi: 10.1177/001872088903100503

Haouij, N. E., Poggi, J. M., Sevestre-Ghalila, S., Ghozi, R., and Jaïdane, M. (2018). "AffectiveROAD system and database to assess driver's attention," in *Proceedings of the 33rd Annual ACM Symposium on Applied Computing* (pp. 800–803).

Hartley, L. R., and El Hassani, J. (1994). Stress, violations and accidents. *Appl. Ergon.* 25, 221–230. doi: 10.1016/0003-6870(94)90003-5

Healey, J. A., and Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intell. Transport. Syst.* 6, 156–166. doi: 10.1109/TITS.2005.848368

Heikoop, D. D., de Winter, J. C., van Arem, B., and Stanton, N. A. (2017). Effects of platooning on signal-detection performance, workload, and stress: a driving simulator study. *Appl. Ergon.* 60, 116–127. doi: 10.1016/j.apergo.2016.10.016

Heikoop, D. D., de Winter, J. C., van Arem, B., and Stanton, N. A. (2019). Acclimatizing to automation: Driver workload and stress during partially automated car following in real traffic. *Transport. Res. part F: Traffic Psychol. Behav.* 65, 503–517. doi: 10.1016/j.trf.2019.07.024

Higgins, J. P. T., Deeks, J. J., and Altman, D. G. (2011). "Chapter 16: Special topics in statistics," in *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011),* eds Higgins JPT, Green S (The Cochrane Collaboration). Available online at: www.handbook.cochrane.org.

Hill, J. D., and Boyle, L. N. (2007). Driver stress as influenced by driving maneuvers and roadway conditions. *Transport. Res. part F: Traffic Psychol. Behav.* 10, 177–186. doi: 10.1016/j.trf.2006.09.002

Hughes, A. M., Hancock, G. M., Marlow, S. L., Stowers, K., and Salas, E. (2019). Cardiac measures of cognitive workload: a meta-analysis. *Hum. Factors* 61, 393–414. doi: 10.1177/0018720819830553

Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Admin. Sci. Q.* 24, 602–611. doi: 10.2307/2392366

Johnson, M. J., Chahal, T., Stinchcombe, A., Mullen, N., Weaver, B., and Bédard, M. (2011). Physiological responses to simulated and on-road driving. *Int. J. Psychophysiol.* 81, 203–208. doi: 10.1016/j.ijpsycho.2011.06.012

Khattak, Z. H., Fontaine, M. D., and Boateng, R. A. (2018). Evaluating the impact of adaptive signal control technology on driver stress and behavior using real-world experimental data. *Transport. Res. part F: Traffic Psychol. Behav.* 58, 133–144. doi: 10.1016/j.trf.2018.06.006

Kim, H. G., Cheon, E. J., Bai, D. S., Lee, Y. H., and Koo, B. H. (2018). Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry Investig.* 15:235. doi: 10.30773/pi.2017.08.17

Kim, S., Rhee, W., Choi, D., Jang, Y. J., and Yoon, Y. (2018). Characterizing driver stress using physiological and operational data from real-world electric vehicle driving experiment. *Int. J. Automot. Technol.* 19, 895–906. doi: 10.1007/s12239-018-0086-0

Kloimüller, I., Karazman, R., Geissler, H., Karazman-Morawetz, I., and Haupt, H. (2000). The relation of age, work ability index and stress-inducing factors among bus drivers. *Int. J. Ind. Ergon.* 25, 497–502. doi: 10.1016/S0169-8141(99)00035-9

Kontogiannis, T. (2006). Patterns of driver stress and coping strategies in a Greek sample and their relationship to aberrant behaviors and traffic accidents. *Accid. Anal. Prevent.* 38, 913–924. doi: 10.1016/j.aap.2006.03.002

Lanatà, A., Valenza, G., Greco, A., Gentili, C., Bartolozzi, R., Bucchi, F., et al. (2014). How the autonomic nervous system and driving style change with incremental stressing conditions during simulated driving. *IEEE Trans. Intell. Transport. Syst.* 16, 1505–1517. doi: 10.1109/TITS.2014.2365681

Langford, C., and Glendon, A. I. (2002). Effects of neuroticism, extraversion, circadian type and age on reported driver stress. *Work Stress* 16, 316–334. doi: 10.1080/0267837021000059019

Lee, D. S., Chong, T. W., and Lee, B. G. (2017). Stress events detection of driver by wearable glove system. *IEEE Sens. J.* 17, 194–204. doi: 10.1109/JSEN.2016.2625323

Lee, H. B., Choi, J. M., Kim, J. S., Kim, Y. S., Baek, H. J., Ryu, M. S., et al. (2007). "Nonintrusive biosignal measurement system in a vehicle," In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (New York, NY: IEEE), 2303–2306.

Li, J., Zhao, X., Xu, S., Ma, J., and Rong, J. (2013). The study of driving simulator validation for physiological signal measures. *Procedia-Soc. Behav. Sci.* 96, 2572–2583. doi: 10.1016/j.sbspro.2013.08.288

Lohani, M., Payne, B. R., and Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Front. Hum. Neurosci.* 13:57. doi: 10.3389/fnhum.2019.00057

Magana, V. C., and Munoz-Organero, M. (2017). Toward safer highways: predicting driver stress in varying conditions on habitual routes. *IEEE Vehicul. Technol. Magazine* 12, 69–76. doi: 10.1109/MVT.2017.2692059

Manseer, M., and Riener, A. (2014). "Evaluation of driver stress while transiting road tunnels," in *Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 1–6).

Matthews, G. (2002). Towards a transactional ergonomics for driver stress and fatigue. *Theoretical Issues Ergon. Sci.* 3, 195–211. doi: 10.1080/14639220210124120

Matthews, G., Desmond, P. A., Joyner, L., Carcary, B., and Kirby, G. (1997). "A comprehensive questionnaire measure of driver stress and affect," in *Traffic and transport psychology. Theory and application*.

Matthews, G., Dorn, L., and Glendon, A. I. (1991). Personality correlates of driver stress. *Pers. Individ. Dif.* 12, 535–549. doi: 10.1016/0191-8869(91)90248-A

Matthews, G., Dorn, L., Hoyes, T. W., Davies, D. R., Glendon, A. I., and Taylor, R. G. (1998). Driver stress and performance on a driving simulator. *Hum. Fact. J. Hum. Fact. Ergonom. Soc.* 40, 136–149. doi: 10.1518/001872098779480569

Matthews, G., Lin, J., and Wohleber, R. (2017). Personality, stress and resilience: a multifactorial cognitive science perspective. *Psihologijske teme* 26, 139–162. doi: 10.31820/pt.26.1.6

Matthews, G., Reinerman-Jones, L. E., Barber, D. J., and Abich, I. V., J. (2015). The psychometrics of mental workload: multiple measures are sensitive but divergent. *Hum. Factors* 57, 125–143. doi: 10.1177/0018720814539505

McCrae, R. R. (1990). Controlling neuroticism in the measurement of stress. *Stress Med.* 6, 237–241. doi: 10.1002/smi.2460060309

Meesit, R., Kanitpong, K., and Jiwattanakulpaisarn, P. (2020). Investigating the influence of highway median design on driver stress. *Transport. Res. Interdisciplin. Perspect.* 4:100098. doi: 10.1016/j.trip.2020.100098

Mehler, B., Reimer, B., and Coughlin, J. F. (2012). Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: an on-road study across three age groups. *Hum. Factors* 54, 396–412. doi: 10.1177/0018720812442086

Miller, E. E., and Boyle, L. N. (2013). Variations in road conditions on driver stress: insights from an on-road study. *Proceed. Hum. Fact. Ergonom. Soc. Ann. Meet.* 57, 1864–1868. doi: 10.1177/1541931213571416

Miller, E. E., and Boyle, L. N. (2015). Driver behavior in road tunnels: association with driver stress and performance. *Transp. Res. Rec.* 2518, 60–67. doi: 10.3141/2518-08

Milleville-Pennel, I., and Charron, C. (2015). Driving for real or on a fixed-base simulator: is it so different? an explorative study. *Presence* 24, 74–91. doi: 10.1162/PRES_a_00216

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 6:e1000097. doi: 10.1371/journal.pmed.1000097

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., and Prisma Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* 4:1. doi: 10.1186/2046-4053-4-1

Morris, D. M., Erno, J. M., and Pilcher, J. J. (2017). Electrodermal response and automation trust during simulated self-driving car use. *Proceed. Hum. Fact. Ergonom. Soc. Ann. Meet.* 61, 1759–1762. doi: 10.1177/1541931213601921

Napoletano, P., and Rossi, S. (2018). "Combining heart and breathing rate for car driver stress recognition," in *2018 IEEE 8th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)* (New York, NY: IEEE), 1–5.

O'Donnell, R. D., and Eggemeier, F. T. (1986). "Workload assessment methodology," in *Handbook of Perception and Human Performance. Cognitive Processes and Performance*, eds Boff, K. R. and J. P. Thomas J. P., Vol. 2., 2nd ed. (Oxford, UK: Wiley), pp. 1–49.

Ollander, S., Godin, C., Charbonnier, S., and Campagne, A. (2016). Feature and sensor selection for detection of driver stress. *PhyCS.* 16, 115–122. doi: 10.5220/0005973901150122

Ooi, J. S. K., Ahmad, S. A., Ishak, A. J., Minhad, K. N., Ali, S. H. M., and Chong, Y. Z. (2018). Grove: an auxiliary device for sympathetic assessment *via* EDA measurement of neutral, stress, and anger emotions during simulated driving conditions. *Int. J. Med. Eng. Inform.* 10, 16–29. doi: 10.1504/IJMEI.2018.090076

Paredes, P. E., Ordonez, F., Ju, W., and Landay, J. A. (2018). "Fast and furious: detecting stress with a car steering wheel," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–12.

Parent, M., Peysakhovich, V., Mandrick, K., Tremblay, S., and Causse, M. (2019). The diagnosticity of psychophysiological signatures: can we disentangle mental workload from acute stress with ECG and fNIRS?. *Int. J. Psychophysiol.* 146, 139–147. doi: 10.1016/j.ijpsycho.2019.09.005

Pedrotti, M., Mirzaei, M. A., Tedesco, A., Chardonnet, J. R., Mérienne, F., Benedetto, S., et al. (2014). Automatic stress classification with pupil diameter analysis. *Int. J. Hum. Comput. Interact.* 30, 220–236. doi: 10.1080/10447318.2013.848320

Pesle, F., Lespinet-Najib, V., Garcia, L., Bougard, C., Diaz, E., and Schneider, S. (2018). ≪ Profils psychologiques et styles de conduite: Analyse exploratoire multidimensionnelle de la vulnérabilité au stress des conducteurs, ≫ in *ERGO'IA 2018*. hal-01882632.

Plarre, K., Raij, A., Hossain, S. M., Ali, A. A., Nakajima, M., Al'Absi, M., et al. (2011). ≪ Continuous inference of psychological stress from sensory measurements collected in the natural environment," in *Proceedings of the 10th ACM/IEEE international conference on information processing in sensor networks* (New York, NY: IEEE), 97–108.

Rastgoo, M. N., Nakisa, B., Maire, F., Rakotonirainy, A., and Chandran, V. (2019). Automatic driver stress level classification using multimodal deep learning. *Expert Syst. Appl.* 138:112793. doi: 10.1016/j.eswa.2019.07.010

Rastgoo, M. N., Nakisa, B., Rakotonirainy, A., Chandran, V., and Tjondronegoro, D. (2018). A critical review of proactive detection of driver stress levels based on multimodal measurements. *ACM Comput. Surv.* 51, 1–35. doi: 10.1145/3186585

Recarte, M. A., and Nunes, L. M. (2003). Mental workload while driving: effects on visual search, discrimination, and decision making. *J. Experiment. Psychol. Appl.* 9:119. doi: 10.1037/1076-898X.9.2.119

Reimer, B., Mehler, B., and Coughlin, J. F. (2016). Reductions in self-reported stress and anticipatory heart rate with the use of a semi-automated parallel parking system. *Appl. Ergon.* 52, 120–127. doi: 10.1016/j.apergo.2015.07.008

Rendon-Velez, E., Van Leeuwen, P. M., Happee, R., Horváth, I., van der Vegte, W. F., and De Winter, J. C. F. (2016). The effects of time pressure on driver performance and physiological activity: a driving simulator study. *Transport. Res. part F: Traffic Psychol. Behav.* 41, 150–169. doi: 10.1016/j.trf.2016.06.013

Rigas, G., Goletsis, Y., and Fotiadis, D. (2012). Real-time driver's stress event detection. *IEEE Transact. Intell. Transport. Syst.* 13, 221–234. doi: 10.1109/TITS.2011.2168215

Rodrigues, J. G., Kaiseler, M., Aguiar, A., Cunha, J. P. S., and Barros, J. (2015). A mobile sensing approach to stress detection and memory activation for public bus drivers. *IEEE Transact. Intell. Transport. Syst.* 16, 3294–3303. doi: 10.1109/TITS.2015.2445314

Roelofs, K. (2017). Freeze for action: neurobiological mechanisms in animal and human freezing. *Philosophic. Transact. Roy. Soc. B: Biol. Sci.* 372:20160206. doi: 10.1098/rstb.2016.0206

Rohatgi, A. (2014). *WebPlotDigitizer user manual version 3.4.* Available online at: https://automeris.io/WebPlotDigitizer/userManual.pdf 1–18.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychol. Bull.* 86, 638–641. doi: 10.1037/0033-2909.86.3.638

Schießl, C. (2008). Subjective strain estimation depending on driving manoeuvres and traffic situation. *IET Intell. Transport Syst.* 2, 258–265. doi: 10.1049/iet-its:20080024

Schmidt-Daffy, M. (2013). Fear and anxiety while driving: Differential impact of task demands, speed and motivation. *Transport. Res. part F: Traffic Psychol. Behav.* 16, 14–28. doi: 10.1016/j.trf.2012.07.002

Seeman, T. E., Singer, B., and Charpentier, P. (1995). Gender differences in patterns of HPA axis response to challenge: Macarthur studies of successful aging. *Psychoneuroendocrinology* 20, 711–725. doi: 10.1016/0306-4530(95)00024-0

Selye, H. (1956). *The Stress of Life*. New York, NY: McGraw-Hill.

Selye, H. (1976). "Stress without distress," in *Psychopathology of Human Adaptation,* ed G. Serban (Boston, MA: Springer).

Simon, F., and Corbett, C. (1996). Road traffic offending, stress, age, and accident history among male and female drivers. *Ergonomics* 39, 757–780. doi: 10.1080/00140139608964497

Singh, R. R., Conjeti, S., and Banerjee, R. (2011). "An approach for real-time stress-trend detection using physiological signals in wearable computing systems for automotive drivers," in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (New York, NY: IEEE), 1477–1482.

Stanton, N. A., and Young, M. S. (1998). Vehicle automation and driving performance. *Ergonomics* 41, 1014–1028. doi: 10.1080/001401398186568

Stanton, N. A., and Young, M. S. (2005). Driver behaviour with adaptive cruise control. *Ergonomics* 48, 1294–1313. doi: 10.1080/00140130500252990

Stephens, C. L., Christie, I. C., and Friedman, B. H. (2010). Autonomic specificity of basic emotions: Evidence from pattern classification and cluster analysis. *Biol. Psychol.* 84, 463–473. doi: 10.1016/j.biopsycho.2010.03.014

Tavakoli, A., Boukhechba, M., and Heydarian, A. (2020). "Personalized driver state profiles: a naturalistic data-driven study," In *International Conference on Applied Human Factors and Ergonomics* (Cham: Springer), 32–39.

Thayer, J. F., Friedman, B. H., and Borkovec, T. D. (1996). Autonomic characteristics of generalized anxiety disorder and worry. *Biol. Psychiatry* 39, 255–266. doi: 10.1016/0006-3223(95)00136-0

Westerman, S. J., and Haigney, D. (2000). Individual differences in driver stress, error and violation. *Pers. Individ. Dif.* 29, 981–998. doi: 10.1016/S0191-8869(99)00249-4

Wickens, C. M., Wiesenthal, D. L., and Roseborough, J. E. (2015). In situ methodology for studying state driver stress: A between-subjects design replication. *J. Appl. Biobehav. Res.* 20, 37–51. doi: 10.1111/jabr.12029

Wohlin, C. (2014). "Guidelines for snowballing in systematic literature studies and a replication in software engineering," In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, 1–10.

Yamakoshi, T., Yamakoshi, K., Tanaka, S., Nogawa, M., Park, S. B., Shibata, M., et al. (2008). "Feasibility study on driver's stress detection from differential skin temperature measurement," in *2008 30th Annual International Conference of* *the IEEE Engineering in Medicine and Biology Society* (New York, NY: IEEE), 1076–1079.

Yamakoshi, T., Yamakoshi, K. I., Tanaka, S., Nogawa, M., Shibata, M., Sawada, Y., et al. (2007). "A preliminary study on driver's stress index using a new method based on differential skin temperature measurement," in 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (New York, NY: IEEE), 722–725.

Young, M. S., and Stanton, N. A. (2002). Attention and automation: new perspectives on mental underload and performance. *Theor. Issues Ergon. Sci.* 3, 178–194. doi: 10.1080/14639220210123789

Yu, Y. J., Yang, Z., Oh, B. S., Yeo, Y. K., Liu, Q., Huang, G. B., et al. (2016)." Investigation on driver stress utilizing ECG signals with on-board navigation systems in use," in *2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV)* (New York, NY: IEEE), 1–6.

Zhao, X., Chi, J., Wu, Y., Rong, J., and Ma, J. (2014). "Impact of chevron alignment sign on driver stress on horizontal curves: a simulator study," in *CICTP 2014: Safe, Smart, and Sustainable Multimodal Transportation Systems*, 2494–2505

Zhao, Y., Yamamoto, T., and Kanamori, R. (2020). Study of older male drivers' driving stress compared with that of young male drivers. *J. Traff. Transport. Eng.* 7, 467–481. doi: 10.1016/j.jtte.2018.10.011

Zontone, P., Affanni, A., Bernardini, R., Del Linz, L., Piras, A., and Rinaldo, R. (2020). Stress evaluation in simulated autonomous and manual driving through the analysis of skin potential response and electrocardiogram signals. *Sensors.* 20:2494. doi: 10.3390/s20092494

Zontone, P., Affanni, A., Bernardini, R., Del Linz, L., Piras, A., and Rinaldo, R. (2021). ≪ Emotional response analysis using electrodermal activity, electrocardiogram and eye tracking signals in drivers with various car setups," in *2020 28th European Signal Processing Conference (EUSIPCO)* (New York, NY: IEEE), 1160–1164.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroergonomics

REVIEW
published: 23 December 2021
doi: 10.3389/fnrgo.2021.765322

The Origins of Passive, Active, and Sleep-Related Fatigue

Steven D. Chong and Carryl L. Baldwin*

Department of Psychology, Program of Human Factors Psychology, Wichita State University, Wichita, KS, United States

Driving is a safety-critical task that requires an alert and vigilant driver. Most research on the topic of vigilance has focused on its proximate causes, namely low arousal and resource expenditure. The present article aims to build upon previous work by discussing the ultimate causes, or the processes that tend to precede low arousal and resource expenditure. The authors review different aspects of fatigue that contribute to a loss of vigilance and how they tend to occur; specifically, the neurochemistry of passive fatigue, the electrophysiology of active fatigue, and the chronobiology of sleep-related fatigue.

Keywords: fatigue, arousal, circadian rhythm, vigilance, driving, automation, attention, vigilance decrement

OPEN ACCESS

Edited by:
Christer Ahlström, Swedish National Road and Transport Research Institute, Sweden

Reviewed by:
Fran Pilkington-Cheney, Loughborough University, United Kingdom
Hadas Chassidim, Sami Shamoon College of Engineering, Israel

****Correspondence:***
Steven D. Chong
stevenchong330@gmail.com

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: *26 August 2021*
Accepted: *03 December 2021*
Published: *23 December 2021*

Citation:
Chong SD and Baldwin CL (2021) The Origins of Passive, Active, and Sleep-Related Fatigue. Front. Neuroergon. 2:765322. doi: 10.3389/fnrgo.2021.765322

INTRODUCTION

According to the Centers for Disease Control Prevention (2020), motor vehicle crashes are the second leading cause of accidental or unintentional death in the United States. Driver distraction and drowsiness are thought to be the most common causes for roadside crashes (AAA Foundation for Traffic Safety, 2018). Driving automation systems (DASs; e.g., adaptive cruise control and active lane keeping) have been introduced to mitigate these crash rates by reducing workload on the driver (Wickens et al., 2010; Wickens, 2018). They may, however, inadvertently introduce new problems to the driver (Mueller et al., 2021), such as increasing the prevalence of driver distraction (Young, 2012; Greenlee et al., 2018), drowsiness (Gimeno et al., 2006; Gaspar et al., 2017; Sikander and Anwar, 2018; Kundinger et al., 2020), engagement in non-driving related tasks (Seppelt and Victor, 2016; Cabrall et al., 2019; Mueller et al., 2021), and loss of situational awareness (Berberian et al., 2017; Brandenburg and Chuang, 2019; Lohani et al., 2019). Overestimating the capabilities of DASs can also lead to overreliance and complacency in the system as well, which may further exacerbate the aforementioned consequences (Parasuraman et al., 2000; Schaefer et al., 2016; Seppelt and Victor, 2016; Hecht et al., 2018). Detailed examination into the processes underlying distraction and sleepiness—or fatigue, more broadly—will be critical to maintaining roadway safety in this time of increasing prevalence of DASs. Here, the authors define fatigue as an adaptive state of stress that occurs due to the interaction between an individual and their environment (Hancock and Warm, 1989). The goal of the present article is to build upon the work of May and Baldwin (2009) by providing more in-depth information on the causality of fatigue, given the findings of newer research. We begin our examination by first discussing a common indicator of fatigue, the vigilance decrement. Then, the remainder of the article will discuss the underlying mechanisms of three dimensions of fatigue: passive, active, and sleep-related fatigue.

The authors want to quickly note that driving using automation will be the primary example used throughout the present paper, however the underlying mechanisms discussed below may also be applicable to other tasks and domains that involve vigilance as well, such as radar operators, anesthesia monitors, air traffic controllers, and cockpit pilots (Donald, 2008; Wiggins, 2011).

THE VIGILANCE DECREMENT

Berberian et al. (2017) identified the vigilance decrement as one of the critical factors that impacts a driver's situational awareness when using automation technologies. The vigilance decrement can be defined as "[t]he deterioration in human performance resulting from adverse working conditions…" (Mackworth, 1948, p. 6). Researchers have studied various aspects that contribute to the aforementioned "adverse working conditions," such as prolonged time-on-task (Mackworth, 1948; Langner and Eickhoff, 2013; Thomson et al., 2015), under-stimulating task features (Scerbo, 1998; Greenlee et al., 2018; Luna et al., 2021), degrading goal maintenance (Hockey, 2011; Braver, 2012; Grahn and Manly, 2012), low critical-event rate (Parasuraman et al., 1987; Langner and Eickhoff, 2013), poor display usability (Hancock, 2013, 2017), and the need to maintain high workload (Helton and Russell, 2017). During a vigilance decrement, one's cognitive state tends to be vulnerable to both distraction (Aston-Jones and Cohen, 2005; Greenlee et al., 2018) and drowsiness (Dinges, 1995; Thiffault and Bergeron, 2003; de Naurois et al., 2019). In this compromised state, a driver is likely to disengage from attentive driving, or monitoring, and instead engage in non-driving related tasks (Randall et al., 2014; Körber et al., 2015; Greenlee et al., 2018).

Most research on vigilance, at least within the Human Factors domain, has focused on the variables and conditions that tend to correlate or precede the vigilance decrement, such as low arousal (Scerbo, 1998), resource depletion (Warm et al., 1996), and neural activity (Smallwood et al., 2012b). Based on this research, two fundamental theories have emerged: the underload, or mindlessness, theory (Sawin and Scerbo, 1995; Scerbo, 1998; Manly et al., 1999) and the overload, or resource depletion, theory (Warm et al., 1996). The underload theory posits that vigilance tasks induce boredom due to their intrinsically monotonous, under-stimulating, and infrequently responding nature. Individuals may, therefore, engage in other activities, such as mind-wandering, in order to elevate their arousal, thereby alleviating boredom (Seli et al., 2015).

The overload theory, on the other hand, suggests that vigilance tasks are inherently difficult due to the high workload associated with having to maintain attention over a prolonged period of time. Specifically, cognitive resources are thought to be allocated toward maintaining proper executive control during a vigil, and once those resources have been exhausted, performance detriments ensue (Randall et al., 2014). Although some have suggested that the under- and overload theories are antithetical to each other, they may describe different dimensions of fatigue.

This has led researchers to adopt an altered perspective regarding the claims made by both underload and overload theories. For instance, some (e.g., Pattyn et al., 2008; Langner and Eickhoff, 2013) view the former as describing the vigilance decrement from an exogenous, "bottom-up" perspective (e.g., monotonous and understimulating task characteristics) and the latter from an endogenous, "top-down" perspective (e.g., subpar executive functioning due to depleted cognitive resources). Similarly, others (e.g., Gimeno et al., 2006; May and Baldwin, 2009; Di Stasi et al., 2012) suggest that these theories describe different aspects of fatigue, passive and active fatigue, respectively. Both underload and overload theories alone, however, only describe the proximate causes of the vigilance decrement, not the underlying, ultimate causes. In other words, low arousal and resource depletion both tend to precede the occurrence of fatigue, but this leads to the question, what precedes low arousal and resource depletion?

PROCESSES UNDERLYING PASSIVE FATIGUE

In the following section, the role of the locus coeruleus and norepinephrine (LC-NE) system will be discussed as it relates to arousal: specifically, the way in which it modulates its firing rate to either broaden or narrow our attentional filter, how it recruits different brain regions to assess the costs and benefits of performing a goal-directed task, and how it works with other brain networks to facilitate exploitative or explorative behavior.

The LC-NE system is one of the primary brainstem neuromodulatory systems that influences arousal (Sara and Bouret, 2012). The locus coeruleus (LC), having broad afferent and efferent connections, is responsible for almost all of the norepinephrine (NE) activity in the neocortex, which regulates the excitatory and inhibitory effects of postsynaptic neurons associated with information selection and processing (Aston-Jones and Cohen, 2005; Bouret and Sara, 2005; Sara and Bouret, 2012). Aston-Jones and Cohen (2005) argued that the LC-NE optimizes reward-contingent behavior by modulating arousal levels such that it either broadens or narrows our attentional filter. The LC-NE either broadens attention in search for a more rewarding task, or it narrows attention to prevent distractions. The LC has two aspects to its firing rate, tonic and phasic. Tonic firing refers to the baseline firing rate and is thought to be indicative of one's current arousal state, while phasic firing refers to the changes in firing in response to task-relevant stimuli presentation and is thought to reflect the degree of cognitive processing (Murphy et al., 2011; Joshi and Gold, 2020). Tonic firing rate, as it relates to performance on a task, resembles the Yerkes-Dodson curve (Yerkes and Dodson, 1908). When the LC exhibits a non-optimal tonic firing rate that is either too high or too low, there are broader neuronal responses to sensory stimuli, which in turn promotes distractibility as it blends the saliency of task-relevant and task-irrelevant stimuli (Mittner et al., 2014, 2016). When the LC exhibits a moderate tonic firing rate, the signal-to noise ratio of phasic responses become salient, which in turn facilitates the discrimination between relevant and irrelevant stimuli, thus promoting sustained attention toward the primary task (Bouret and Sara, 2005).

The firing rate of the LC-NE system has been commonly studied by observing changes in pupillometry due to the strong correlations among LC-NE activity, pupillary dynamics, and arousal state (Mittner et al., 2014, 2016; Hopstaken et al., 2015; Lohani et al., 2019; McWilliams and Ward, 2021). For instance, in a series of experiments, Unsworth and Robison (2018) found that smaller pretrial pupil sizes (as an index of tonic firing) and smaller task-evoked pupillary responses (as an index of phasic firing) were related to lower arousal, poorer performance, and more task-unrelated thoughts. Similarly, Körber et al. (2015) found that passive fatigue 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, greater reports of mind-wandering at the end of the monitoring task, and a general trend of slower reaction times on an auditory oddball task. Although the functional and anatomical mechanisms are not yet fully understood [but see Joshi and Gold (2020), Murphy et al. (2014)], previous research suggests that a task's utility—or the costs and benefits associated with performance on a task—may dictate the firing rate of the LC.

The anterior cingulate cortex (ACC) and the orbitofrontal cortex (OFC) project information regarding a task's utility to the LC-NE system (Aston-Jones and Cohen, 2005). As the costs associated with a task increase—as evaluated by the ACC—the LC exhibits high tonic firing for the search of alternative forms of reward, and it promotes engagement in other tasks (e.g., mind-wandering). On the other hand, as the benefits outweigh the costs—as evaluated by the OFC—the LC exhibits moderate tonic firing, which in turn helps prevent attention from being oriented to task-irrelevant stimuli. The ACC and OFC, receiving inputs from various somatosensory and limbic structures, evaluate a task's utility both in the short- and long-term.

During initial engagement of a goal-directed task, reinforcement learning occurs to favor reward-contingent behavior, and when the ACC detects non-conducive behavioral deviations (e.g., lapses in attention) the prefrontal cortex is recruited to exercise top-down control to calibrate behavior, by way of increasing LC phasic activity, to maintain optimal performance (Aston-Jones and Cohen, 2005; Sara and Bouret, 2012; Massar et al., 2016; Bier et al., 2019). As engagement in the task continues, the costs associated with optimal performance tend to increase, while the rewards associated with the task tend to decrease due to increased satiety, exposure, or predictability of the rewards, eventually resulting in increased LC tonic activity.

For instance, Massar et al. (2016) found that performance-contingent rewards fostered better performance and longer engagement in a vigilance task. Moreover, pretrial pupil size was also greater for those who received rewards as well, suggesting greater engagement in the task. Similarly, when drivers interacted with a system that gamified driving and rewarded drivers for good performance (e.g., degree of lateral steering control, hazard avoidance, etc.) during a manual drive, subjective fatigue was delayed, standard deviation of lane position and unintentional lane crossings were reduced, drivers were less prone to accidents, and there was better compliance with driving at the speed limit (Bier et al., 2019). Rewards may be introduced to help rebalance a task's utility by fostering motivation and optimizing one's arousal state for the task at hand (Aston-Jones and Cohen, 2005; Boksem and Tops, 2008). Some have even suggested that NE acts as a "network resetting" (Bouret and Sara, 2005) or "circuit-breaking" (Corbetta and Shulman, 2002) signal that reconfigures network connectivity for the purpose of maximizing rewarding outcomes.

The LC-NE system may influence broad network connectivity possibly due to its connections to the frontoparietal control network (FPCN; Sara and Bouret, 2012). According to the global workspace theory, the integration of various neural submodules forms conscious experience and underlies the engagement in various cognitive processes (Baars et al., 2003; Smallwood et al., 2012a). The FPCN is thought to be responsible for housing a "global workspace" that consolidates all cortical communication and facilitates the activity of the most salient submodule. Because of this, the FPCN plays a critical role in not only directing attention, for instance, based on the most dominant submodule, but also maintaining attention as well, be it externally or internally oriented.

Within a vigilance context, two key networks that contribute to the global workspace of the FPCN are the dorsal attention network (DAN) and the default mode network (DMN; Dang et al., 2012). When attention is oriented externally, the FPCN couples with the DAN, while coupling with the DMN when attention is directed internally (Spreng, 2012). The LC-NE may be one of the primary neuromodulatory systems, alongside the ventral tegmental area-dopamine system (VTA-DA), that dictates which network the FPCN couples with by adjusting the gain of neuronal activity (Dang et al., 2012; Ranjbar-Slamloo and Fazlali, 2020).

Neural gain signifies the specificity of functional connectivity and a greater signal-to-noise ratio of strong neuronal communication (Mittner et al., 2016). For instance, when neural gain is low, broad functional connectivity can be observed, with weaker connections becoming more active and therefore competitive with stronger connections. This would translate to greater distractibility and engagement in other tasks (e.g., mind wandering). Conversely, during high neural gain, functional connectivity becomes precise by suppressing weaker connections, while stronger, task-relevant connections remain dominant. While in a state of low arousal, tonic LC activity tends to be high, which in turn reduces neural gain and allows task-unrelated cortical activity (e.g., DMN) to influence the global workspace of the FPCN to facilitate explorative behavior. In contrast, while in a state of optimal arousal, phasic LC activity tends to be high, which increases neural gain, and promotes exploitative behavior by facilitating the coupling of the FPCN and the DAN (Aston-Jones and Cohen, 2005; Bouret and Sara, 2005; Sara and Bouret, 2012; Mittner et al., 2016).

In sum, the LC-NE system is implicated in a wide range of cognitive processes, such as attention, memory, mood, motivation, perception, and arousal. This is likely due to its broad afferent and efferent connections and the fact that it is the primary source of NE in the neocortex (Aston-Jones and Cohen, 2005; Bouret and Sara, 2005; Sara and Bouret, 2012). The LC-NE, receiving input from the ACC and OFC, plays a critical role in optimizing task performance by modulating arousal levels, which in turn broadens or narrows our attentional filter for rewarding outcomes. Within a vigilance context, it can also influence the FPCN by adjusting the neural gain associated with environmental stimuli, thereby facilitating the coupling between the DAN (task-related) or the DMN (task-unrelated). When a driver is tasked with monitoring the automation system, the understimulating nature of the task may induce hypo-arousal by influencing the firing rate of the LC-NE due to an imbalance of the costs for having to maintain vigilance (Körber et al., 2015). Additionally, this imbalance may foster the activity of task-unrelated brain networks and promote engagement in non-driving related tasks such as mind-wandering (Bier et al., 2019). The LC-NE system is a critical component in understanding the cause of low arousal, or passive fatigue. Passive fatigue, however, only addresses one aspect of fatigue; depleted cognitive resources, or active fatigue, can address another.

PROCESSES UNDERLYING ACTIVE FATIGUE

Active fatigue differentiates itself from passive fatigue by examining fatigue as a function of cognitive load and time on task (Warm et al., 1996; Grier et al., 2003; Szalma and Claypoole, 2019), as opposed to arousal (e.g., Scerbo, 1998; Hockey, 2011; Langner and Eickhoff, 2013). Most of the explanation regarding the link between active fatigue and its two components are based upon "cognitive resources;" however, despite decades of research it has yet to be objectively, or even collectively, defined (Dehais et al., 2020). Below, the present authors build upon those previous works by describing active fatigue from a more objective perspective—as a function of long-term neuronal potentiation.

It is generally thought that the act of maintaining vigilance over a prolonged period of time is very taxing on cognitive resources (Warm et al., 1996). This idea intuitively makes sense, given that some sort of "cost" is always associated with some sort of activity in the real world (e.g., time, attention, money, etc.). However, "cognitive resources" have remained obtuse and no clear definition has been presented (Dehais et al., 2020). The notion of some "cost" or resource requirement for any cognitive activity, however, should not wholly be discarded. Specifically, in the context of vigilance, a decrement could occur not due to an over-expenditure of cognitive resources *per se*, but rather an overaccumulation of synaptic load. First, however, it is important to clarify the meaning of resource depletion by disentangling correlates of brain activity from indicators of brain metabolism.

Brain activation refers to the changes in blood flow, specifically the arterial oxygen concentration, as reflected by the oxygen extraction fraction (OEF) when using positron emission tomography (PET) or the blood-oxygen-level-dependent (BOLD) signal when using functional magnetic resonance imaging (fMRI). Brain metabolism, on the other hand, refers to the oxidation of glucose, through aerobic glycolysis and oxidative phosphorylation, to create adenosine triphosphate (ATP; Raichle and Gusnard, 2002; Raichle and Mintun, 2006). Simply put, the former refers to oxygen delivery, while the latter refers to oxygen consumption. When researchers observe "brain activity," oxygen delivery is increased in a local region of the brain, under the assumption that greater neuronal activity occurred in that region; oxygen consumption. However, although it does increase slightly, it does not increase to the same degree (Raichle and Gusnard, 2002). Changes in oxygen delivery in a specific region of the brain does not necessarily entail meaningful changes in oxygen consumption, or energy expenditure. In other words, brain activation and brain metabolism are two distinct processes. They are related, yet independent.

One of the assumptions underlying BOLD signals, for instance, is that local blood-flow changes supply the necessary ingredients (oxygen and glucose) to create ATP to fuel task-induced brain activity. Raichle and Mintun (2006), however, argued that this assumption is somewhat misguided, if not wholly incorrect. They note that the genesis of this assumption stems from research showing the relatively strong correlations of single-unit recordings, multiunit recordings, and local field potentials with changes in fMRI BOLD signals. However, this only demonstrates a correlational relationship between local blood-flow changes and neuronal activity - it does not demonstrate a causal relationship. Moreover, single-unit and multiunit recordings represent different aspects of neuronal activity compared to local field potentials. Specifically, the former refers to the spiking activity of neurons, or the output, while the latter refers to the membrane currents of neurons, or the input. Therefore, it is unclear why and how changes in local blood-flow are related to neuronal activity. In addition, there is usually a lag time of 4–6 s regarding task-induced changes in BOLD signals. The brain, as Raichle and Mintun argued, would not depend on such a slow process to provide the necessary moment-to-moment prerequisites for brain activity. Instead, it would be more efficient to extract more of the oxygen that is already circulating in the blood and to use the glucose that is already stored in the glycogen of astrocytes, suggesting that the brain does not necessarily depend on increases in local blood-flow to fuel brain activity.

Interestingly, there are no significant changes in whole-brain blood flow due to the engagement of a goal-directed task, with only negligible ($\leq$5%) differences in local blood-flow (Raichle and Gusnard, 2002). Moreover, the brain's metabolism, or energy budget, also remains strikingly constant as well, again irrespective of engaging in a task or passively at rest, suggesting that the traditional paradigm of resource depletion may not be the most appropriate conceptualization of active fatigue. When taking into account that the vast majority of the brain's metabolism is allocated toward maintaining proper excitatory and inhibitory synaptic activity (Raichle and Mintun, 2006), investigating synaptic processes, specifically its homeostasis, could help illuminate the fatiguing nature of vigilance tasks.

Every animal—be it terrestrial, oceanic, or avian—engages in sleep to regulate synaptic activity that took place during wakefulness (Vyazovskiy and Harris, 2013). With prolonged wakefulness, cognitive and physical deficits, and even death, inevitably occur (Wang et al., 2011). The synaptic homeostasis hypothesis (Tononi and Cirelli, 2006) can help explain the rebalancing that takes place as a function of both wakefulness and sleep, and it makes four claims regarding how the brain achieves equilibrium: (1) synaptic potentiation occurs predominantly during active wakefulness. During wakefulness, plastic changes, specifically long-term potentiation, occur due to the presynaptic firing and postsynaptic depolarization associated with a broad range of neuronal activity. Evidence comes from the increases in synaptic density, due to long-term potentiation, that are typically found when animals are in stimulus-enriched environments.

(2) Slow wave activity (SWA) regulates the homeostasis of synaptic potentiation. Predominantly observed in non-rapid eye movement (NREM) sleep, SWA are spontaneous oscillations consisting of low frequency (<1 Hz), high amplitude (>140 μv) sequences of synchronized depolarized up-phases (On-periods) and hyperpolarized down-phases (Off-periods; Van Someren et al., 2011). Delta waves (1–4 Hz) with amplitudes ranging from 75 to 140 μv have also been thought to reflect slow wave activity as well but with less cortical synchronization. A localized concentration of SWA is commonly observed in a location-dependent manner based on the area of synaptic potentiation that occurred during wake in both cortical and subcortical regions (Vyazovskiy and Harris, 2013; Bernardi et al., 2015; Quercia et al., 2018; Andrillon et al., 2019). (3) Slow wave activity facilitates homeostasis primarily through synaptic downscaling. Synaptic weight, or load, is accumulated onto neurons as a function of use-dependent, long-term potentiation. Slow wave activity acts to proportionally downscale, or decrease, the synaptic weight of neurons engaged in long-term potentiation, thereby resetting the net load accumulated during wake. Specifically, the magnitude of SWA tends to correlate with how long one prolongs wakefulness and exponentially decreases as one remains asleep (Vyazovskiy et al., 2011). Finally, (4) synaptic downscaling is one of the ultimate causes of cognitive restoration. Evidence for this claim comes from the uncompromising, and intrusive, occurrences of SWA during prolonged wakefulness, such as micro-sleeps (Dinges, 1995). Moreover, irregularities in SWA have been connected to various mental disorders, such as depression, and sleep-related disorders, such as insomnia (Tononi and Cirelli, 2006). In fact, Vyazovskiy and Harris (2013) suggest not only does the resetting of neuronal firing rates plays a critical role in cognitive restoration, but that SWA may be a self-defense countermeasure that acts as cellular maintenance to prevent unnecessary, long-term damage (e.g., excessive oxidative stress and damage to DNA, proteins, and lipids). For instance, in the face of cognitively demanding tasks, neurons will maintain high synaptic activity for optimal performance until neuronal fatigue sets in, as indicated by periods of neuronal silencing. As neurons continue to fatigue, SWA becomes more intrusive in terms of frequency and spatial location, translating into greater decrements in performance (Van Someren et al., 2011; Andrillon et al., 2019).

To summarize, active fatigue may not be best described as an expenditure of finite cognitive resources *per se*, because the brain's energy budget remains relatively constant, irrespective of task engagement (Raichle and Gusnard, 2002; Raichle and Mintun, 2006). Instead, we propose that active fatigue may be more accurately thought of as long-term potentiation of neurons inducing a high synaptic load (Tononi and Cirelli, 2006). Evidence for this stems from the observation that the vast majority of the brain's energy budget is allocated toward regulating excitatory and inhibitory synaptic activity. When neuronal fatigue occurs, the brain will engage in sleep-related processes, such as SWA, to compensate for high synaptic load (Vyazovskiy and Harris, 2013). For instance, Bernardi et al. (2015) demonstrated indicators of SWA in visuomotor and executive functioning areas when participants were controlling a driving simulator for a prolonged time, and the presence of SWA was associated with poorer performance. Sleep-related processes during cognitively demanding tasks could result in not only poorer performance but also promote engagement in unrelated tasks, such as mind-wandering (Andrillon et al., 2019). Alternating SWA in different neural locations (e.g., DAN vs. DMN), for example, may represent a major source of naturally occurring cognitive restoration and homeostasis.

CIRCADIAN PROCESSES SUBSERVE PASSIVE AND ACTIVE FATIGUE

Independent of task characteristics, arousal can also fluctuate based on our circadian rhythm (Carrier and Monk, 2000; Aston-Jones et al., 2001). Our circadian rhythm, or clock system, is primarily controlled by a central clock—the suprachiasmatic nucleus of the hypothalamus (SCN)—with influences coming from a myriad of peripheral clocks found in all our tissue (Nicolaides et al., 2014). The SCN plays a significant role in our sleep-wake cycle by increasing alertness through arousal regulation (Aston-Jones et al., 2001). Specifically, circadian variations in arousal are influenced by a circuit consisting of the SCN, the dorsomedial nucleus of the hypothalamus (DMH), and the locus coeruleus (LC). Light/dark cues from the environment modulate SCN activity. The SCN sends projections to the DMH, which acts a relay for the LC, and the LC, in turn, elicits a NE response, thereby influencing arousal. The SCN-DMH-LC circuit may partially explain the ultradian time-of-day variations found in a variety of cognitive processes, such as fatigue, alertness, and memory (Carrier and Monk, 2000; Van Dongen and Dinges, 2000; Aston-Jones et al., 2007). Moreover, NE in blood plasma also exhibits a circadian rhythm with numerous ultradian peaks throughout wakefulness (Sowers and Vlachakis, 1984). Interestingly, these NE rhythms also closely align with time-of-day variations found in mind-wandering rates (Smith et al., 2018), subjective alertness, arousal ratings, and driving performance (Lenné et al., 1997). Not only do circadian processes independently influence passive fatigue (Carrier and Monk, 2000; Aston-Jones et al., 2001), they can also independently influence active fatigue as well (Bernardi et al., 2015). Before describing these influences, it is important to acknowledge the blurred boundaries between wakefulness and sleep.

Wakefulness and sleep are generally thought of as two distinct states, but electrophysiological evidence indicates they can occur simultaneously. Wakefulness can generally be thought of as a state in which most of the brain is active, while in a sleep state, most of the brain is quiescent. But, cortical and subcortical regions can briefly engage in sleep-related processes while other regions remain "awake," referred to as local sleep (Vyazovskiy and Harris, 2013). Local sleep is defined as the "transient, regional neurophysiological state showing a mixture of features characteristic of (i) wakefulness and sleep, (ii) different sleep stages (NREM and REM sleep), or (iii) different sleep depths (light or deep sleep)" (Andrillon et al., 2019, p. 2). One indicator of local sleep is the presence of high-amplitude, slow oscillations (i.e., SWA). As previously described, slow oscillations are a series of synchronized neuronal activity (On-periods) and neuronal silencing (Off-periods) among different cortical areas that can occur locally in terms of both time and space (Vyazovskiy et al., 2011).

The occurrence of local sleep, as indexed by SWA, is dependent on two factors: time spent awake and use-dependent activity (Vyazovskiy et al., 2011; Bernardi et al., 2015; Quercia et al., 2018; Andrillon et al., 2019). In terms of time, local sleep tends to be more frequent as sleep pressure, or the need to sleep, builds. Conversely, local sleep tends to be less frequent the longer one remains in sleep (Van Someren et al., 2011). For instance, local sleep rarely occurs in the first few hours after wake, but gradually appears as one remains awake (Vyazovskiy et al., 2011). Microsleeps are an extreme example of local sleep occurring during wakefulness (Dinges, 1995; Andrillon et al., 2019) and have been associated with poorer driving (Boyle et al., 2008) and increased accident rates (Sirois et al., 2009). In NREM sleep, SWA tends to be most prominent in frontal and parietal regions (Vyazovskiy et al., 2011). Though both areas generally tend to exhibit synchronized SWA concurrently (global), they can occasionally—particularly during wakefulness—occur independently (local) as well. As the number of occurrences of SWA increases, more areas of the brain engage in SWA (Andrillon et al., 2019). Drastic changes in various neuromodulators could also be indicators of local sleep as well.

From a circadian perspective, we exhibit three vigilance states: wakefulness, NREM sleep, and REM sleep (Aston-Jones et al., 2001, 2007). These three states are distinguishable based on various electrophysiological, physiological, and neurochemical metrics. During wakefulness, high-frequency, low-amplitude activities are common, specifically in the gamma (30–60 Hz) and beta (20–30 Hz) frequencies. We also generally have higher heart rates and heart rate variability compared to the other vigilance states. During NREM sleep, low-frequency, high-amplitude activities become common, such as frequent occurrences of sleep spindles (12–15 Hz), delta waves (1–4 Hz), and slow waves (<1 Hz). Heart rate and brain temperature also decrease during this NREM state as well, compared to wakefulness. Finally, in REM sleep, we exhibit similar brain patterns as wakefulness (high-frequency, low-amplitude signals). In fact, REM sleep is also referred to as paradoxical sleep due to its striking resemblance to wakefulness. It is impossible to distinguish REM sleep from wakefulness when only observing one's electrophysiology. The differences between REM sleep and wakefulness can be found when examining one's physiology, specifically transient paralysis in the muscles accompanied by rapid changes in body temperature during REM sleep. There are neurochemical differences between the three vigilance states. Significant changes in tonic firing of NE occur such that the firing rate tends to be highest during wakefulness, then decreases dramatically during NREM sleep, and becomes almost completely dormant during REM sleep. Specifically, changes in NE occur prior to the transitions between vigilance states (Aston-Jones et al., 2007).

To summarize, passive and active fatigue are distinct constructs, yet they are related in that they independently operate under a circadian rhythm (Carrier and Monk, 2000; Aston-Jones et al., 2001). The SCN-DMH-LC circuit regulates our arousal based on light/dark cues from the environment, and because of this, arousal fluctuates throughout the day due to our sleep-wake cycle (Aston-Jones et al., 2001). Sleep and wakefulness, however, are not mutually exclusive states. Specifically, cortical regions engage in sleep-related processes (i.e., local sleep) based on the degree of neuronal activity that has occurred in that region during wakefulness (Quercia et al., 2018; Andrillon et al., 2019). In this way, circadian processes sit at the junction between passive and active fatigue, and this triumvirate could explain how and why fatigue occurs.

CONCLUSION

Driver distraction and drowsiness, or fatigue more broadly, have continued to be some of the leading causes for motor vehicle crashes, and these crashes have continued to be one of the leading causes of death in the United States. Understanding the mechanisms of cause-and-effects for why fatigue occurs is critical to improving road safety. Fatigue is a multidimensional construct that may have at least three components: passive, active, and sleep-related fatigue. Although these components have different causes, they are interrelated in that they influence each other. Passive fatigue tends to occur due to various exogenous characteristics inducing hypo-arousal and it can promote engagement in non-driving related tasks to alleviate boredom. Active fatigue occurs as a function of long-term neuronal potentiation in which specific brain regions will engage in sleep-related processes to avoid unnecessary long-term damage due to prolonged activity. Despite these disparate processes, they are both influenced by our circadian rhythm. In other words, time-of-day moderates how passive and active fatigue occurs.

Finally, we would be remiss for not also mentioning alternative theories, outside of the under- and overload theories described above, whose shoulders' this article stands upon. Two, in particular, warrant specific mentioning. First is Hockey's (2011) motivational control theory that provides a compelling case for the dynamic interplay between fatigue and recovery, and second are the works of Hancock (2013, 2017) who elegantly described the weaknesses of previous theories and discusses the impact those weaknesses have had in our understanding of vigilance. The interested reader is highly encouraged to read these original works to challenge their assumptions regarding the latent-variable construct of fatigue.

It is our hope that this review increases general understanding of the specific processes that subserve different aspects of fatigue. And further, that understanding the mechanisms underlying passive and active fatigue, and the influence of circadian rhythms on each will facilitate the development of effective driver fatigue countermeasures for each type.

AUTHOR CONTRIBUTIONS

All authors contributed to the article and approved the submitted version.

REFERENCES

AAA Foundation for Traffic Safety (2018). *2017 Traffic Safety Culture Index*. Washington, DC: AAA Foundation for Traffic Safety. Available online at: https://aaafoundation.org/wp-content/uploads/2018/03/TSCI-2017-Report.pdf (accessed March 11, 2021).

Andrillon, T., Windt, J. M., Silk, T., Drummond, S., Bellgrove, M., and Tsuchiya, N. (2019). Does the mind wander when the brain takes a break? Local sleep in wakefulness, attentional lapses and mind-wandering. *Front. Neurosci*. 13:949. doi: 10.3389/fnins.2019.00949

Aston-Jones, G., Chen, S., Zhu, Y., and Oshinsky, M. L. (2001). A neural circuit for circadian regulation of arousal. *Nat. Neurosci*. 4, 732–738. doi: 10.1038/89522

Aston-Jones, G., and Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci*. 28, 403–450. doi: 10.1146/annurev.neuro.28.061604.135709

Aston-Jones, G., Gonzalez, M., and Doran, S. (2007). "Role of the locus coeruleus-norepinephrine system in arousal and circadian regulation of the sleep-wake cycle," in *Brain Norepinephrine: Neurobiology and Therapeutics,* eds G. A. Ordway, M. A. Schwartz, and A. Frazer (Cambridge: Cambridge University Press), 157–195. doi: 10.1017/CBO9780511544156.007

Baars, B. J., Ramsøy, T. Z., and Laureys, S. (2003). Brain, conscious experience and the observing self. *Trends Neurosci*. 26, 671–675. doi: 10.1016/j.tins.2003.09.015

Berberian, B., Somon, B., Sahaï, A., and Gouraud, J. (2017). The out-of-the-loop Brain: a neuroergonomic approach of the human automation interaction. *Annu. Rev. Contr*. 44, 303–315. doi: 10.1016/j.arcontrol.2017.09.010

Bernardi, G., Siclari, F., Yu, X., Zennig, C., Bellesi, M., Ricciardi, E., et al. (2015). Neural and behavioral correlates of extended training during sleep deprivation in humans: evidence for local, task-specific effects. *J. Neurosci*. 35, 4487–4500. doi: 10.1523/JNEUROSCI.4567-14.2015

Bier, L., Emele, M., Gut, K., Kulenovic, J., Rzany, D., Peter, M., et al. (2019). Preventing the risks of monotony related fatigue while driving through gamification. *Eur. Transport Res. Rev*. 11, 1–19. doi: 10.1186/s12544-019-0382-4

Boksem, M. A., and Tops, M. (2008). Mental fatigue: costs and benefits. *Brain Res. Rev*. 59, 125–139. doi: 10.1016/j.brainresrev.2008.07.001

Bouret, S., and Sara, S. J. (2005). Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. *Trends Neurosci*. 28, 574–582. doi: 10.1016/j.tins.2005.09.002

Boyle, L. N., Tippin, J., Paul, A., and Rizzo, M. (2008). Driver performance in the moments surrounding a microsleep. *Transport. Res. F Traffic Psychol. Behav*. 11, 126–136. doi: 10.1016/j.trf.2007.08.001

Brandenburg, S., and Chuang, L. (2019). Take-over requests during highly automated driving: how should they be presented and under what conditions? *Transport. Res. F Traffic Psychol. Behav*. 66, 214–225. doi: 10.1016/j.trf.2019.08.023

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn. Sci*. 16, 106–113. doi: 10.1016/j.tics.2011.12.010

Cabrall, C. D., Eriksson, A., Dreger, F., Happee, R., and de Winter, J. (2019). How to keep drivers engaged while supervising driving automation? A literature survey and categorisation of six solution areas. *Theoret. Issues Ergon. Sci*. 20, 332–365. doi: 10.1080/1463922X.2018.1528484

Carrier, J., and Monk, T. H. (2000). Circadian rhythms of performance: new trends. *Chronobiol. Inte*. 17, 719–732. doi: 10.1081/CBI-100102108

Centers for Disease Control and Prevention (2020). *Underlying Cause of Death 1999-2019 on CDC WONDER Online Database, Released in 2020. Data Are From the Multiple Cause of Death Files, 1999–2019, as Compiled From Data Provided by the 57 Vital Statistics Jurisdictions Through the Vital Statistics Cooperative Program*. National Center for Health Statistics. Available online at: http://wonder.cdc.gov/ucd-icd10.html (accessed March 24, 2021).

Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci*. 3, 201–215. doi: 10.1038/nrn755

Dang, L. C., O'Neil, J. P., and Jagust, W. J. (2012). Dopamine supports coupling of attention-related networks. *J. Neurosci*. 32, 9582–9587. doi: 10.1523/JNEUROSCI.0909-12.2012

de Naurois, C. J., Bourdin, C., Stratulat, A., Diaz, E., and Vercher, J. L. (2019). Detection and prediction of driver drowsiness using artificial neural network models. *Accid. Anal. Prev*. 126, 95–104. doi: 10.1016/j.aap.2017.11.038

Dehais, F., Lafont, A., Roy, R., and Fairclough, S. (2020). A neuroergonomics approach to mental workload, engagement and human performance. *Front. Neurosci*. 14:268. doi: 10.3389/fnins.2020.00268

Di Stasi, L. L., Renner, R., Catena, A., Cañas, J. J., Velichkovsky, B. M., and Pannasch, S. (2012). Towards a driver fatigue test based on the saccadic main sequence: a partial validation by subjective report data. *Transport. Res. C Emerg. Technol*. 21, 122–133. doi: 10.1016/j.trc.2011.07.002

Dinges, D. F. (1995). An overview of sleepiness and accidents. *J. Sleep Res*. 4, 4–14. doi: 10.1111/j.1365-2869.1995.tb00220.x

Donald, F. M. (2008). The classification of vigilance tasks in the real world. *Ergonomics* 51, 1643–1655. doi: 10.1080/00140130802327219

Gaspar, J. G., Brown, T. L., Schwarz, C. W., Lee, J. D., Kang, J., and Higgins, J. S. (2017). Evaluating driver drowsiness countermeasures. *Traffic Injury Prev*. 18, S58–S63. doi: 10.1080/15389588.2017.1303140

Gimeno, P., Cerezuela, G., and Montanes, M. (2006). On the concept and measurement of driver drowsiness, fatigue and inattention: implications for countermeasures. *Int. J. Vehicle Design* 42, 67–86. doi: 10.1504/IJVD.2006.010178

Grahn, J. A., and Manly, T. (2012). Common neural recruitment across diverse sustained attention tasks. *PLoS ONE* 7:e49556. doi: 10.1371/journal.pone.0049556

Greenlee, E. T., DeLucia, P. R., and Newton, D. C. (2018). Driver vigilance in automated vehicles: hazard detection failures are a matter of time. *Hum. Factors* 60, 465–476. doi: 10.1177/0018720818761711

Grier, R. A., Warm, J. S., Dember, W. N., Matthews, G., Galinsky, T. L., Szalma, J. L., et al. (2003). The vigilance decrement reflects limitations in effortful attention, not mindlessness. *Hum. Factors* 45, 349–359. doi: 10.1518/hfes.45.3.349.27253

Hancock, P. A. (2013). In search of vigilance: the problem of iatrogenically created psychological phenomena. *Am. Psychol*. 68:97. doi: 10.1037/a0030214

Hancock, P. A. (2017). On the nature of vigilance. *Hum. Factors* 59, 35–43. doi: 10.1177/0018720816655240

Hancock, P. A., and Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Hum. Factors* 31, 519–537. doi: 10.1177/001872088903100503

Hecht, T., Feldhütter, A., Radlmayr, J., Nakano, Y., Miki, Y., Henle, C., et al. (2018). "A review of driver state monitoring systems in the context of automated driving," in *Congress of the International Ergonomics Association* (Cham: Springer), 398–408. doi: 10.1007/978-3-319-96074-6_43

Helton, W. S., and Russell, P. N. (2017). Rest is still best: the role of the qualitative and quantitative load of interruptions on vigilance. *Hum. Factors* 59, 91–100. doi: 10.1177/0018720816683509

Hockey, G. R. J. (2011). "A motivational control theory of cognitive fatigue," in *Decade of Behavior/Science Conference. Cognitive Fatigue: Multidisciplinary Perspectives on Current Research and Future Applications*, ed P. L. Ackerman (Washington, DC: American Psychological Association), 167–187. doi: 10.1037/12343-008

Hopstaken, J. F., Van Der Linden, D., Bakker, A. B., and Kompier, M. A. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement. *Psychophysiology* 52, 305–315. doi: 10.1111/psyp.12339

Joshi, S., and Gold, J. I. (2020). Pupil size as a window on neural substrates of cognition. *Trends Cogn. Sci*. 24, 466–480. doi: 10.1016/j.tics.2020.03.005

Körber, M., Cingel, A., Zimmermann, M., and Bengler, K. (2015). Vigilance decrement and passive fatigue caused by monotony in automated driving. *Proc. Manufactur*. 3, 2403–2409. doi: 10.1016/j.promfg.2015.07.499

Kundinger, T., Riener, A., Sofra, N., and Weigl, K. (2020). "Driver drowsiness in automated and manual driving: insights from a test track study," in *Proceedings of the 25th International Conference on Intelligent User Interfaces* (Cagliari), 369–379. doi: 10.1145/3377325.3377506

Langner, R., and Eickhoff, S. B. (2013). Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. *Psychol. Bullet*. 139:870. doi: 10.1037/a0030694

Lenné, M. G., Triggs, T. J., and Redman, J. R. (1997). Time of day variations in driving performance. *Accid. Anal. Prev*. 29, 431–437. doi: 10.1016/S0001-4575(97)00022-5

Lohani, M., Payne, B. R., and Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Front. Hum. Neurosci*. 13:57. doi: 10.3389/fnhum.2019.00057

Luna, F. G., Barttfeld, P., Martín-Arévalo, E., and Lupiáñez, J. (2021). The ANTI-Vea task: analyzing the executive and arousal vigilance decrements while measuring the three attentional networks. *Psicológica* 2021:1. doi: 10.2478/psicolj-2021-0001

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quart. J. Exp. Psychol.* 1, 6–21. doi: 10.1080/17470214808416738

Manly, T., Robertson, I. H., Galloway, M., and Hawkins, K. (1999). The absent mind: further investigations of sustained attention to response. *Neuropsychologia* 37, 661–670. doi: 10.1016/S0028-3932(98)00127-4

Massar, S. A., Lim, J., Sasmita, K., and Chee, M. W. (2016). Rewards boost sustained attention through higher effort: a value-based decision making approach. *Biol. Psychol.* 120, 21–27. doi: 10.1016/j.biopsycho.2016.07.019

May, J. F., and Baldwin, C. L. (2009). Driver fatigue: the importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. *Transport. Res. F Traffic Psychol. Behav*. 12, 218–224. doi: 10.1016/j.trf.2008.11.005

McWilliams, T., and Ward, N. (2021). Underload on the road: measuring vigilance decrements during partially automated driving. *Front. Psychol*. 12:1113. doi: 10.3389/fpsyg.2021.631364

Mittner, M., Boekel, W., Tucker, A. M., Turner, B. M., Heathcote, A., and Forstmann, B. U. (2014). When the brain takes a break: a model-based analysis of mind wandering. *J. Neurosci*. 34, 16286–16295. doi: 10.1523/JNEUROSCI.2062-14.2014

Mittner, M., Hawkins, G. E., Boekel, W., and Forstmann, B. U. (2016). A neural model of mind wandering. *Trends Cogn. Sci*. 20, 570–578. doi: 10.1016/j.tics.2016.06.004

Mueller, A. S., Reagan, I. J., and Cicchino, J. B. (2021). Addressing driver disengagement and proper system use: human factors recommendations for level 2 driving automation design. *J. Cogn. Eng. Decision Mak.* 15, 3–27. doi: 10.1177/1555343420983126

Murphy, P. R., O'connell, R. G., O'sullivan, M., Robertson, I. H., and Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Hum. Brain Map*. 35, 4140–4154. doi: 10.1002/hbm.22466

Murphy, P. R., Robertson, I. H., Balsters, J. H., and O'connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus–noradrenergic arousal function in humans. *Psychophysiology* 48, 1532–1543. doi: 10.1111/j.1469-8986.2011.01226.x

Nicolaides, N. C., Charmandari, E., Chrousos, G. P., and Kino, T. (2014). Circadian endocrine rhythms: the hypothalamic–pituitary–adrenal axis and its actions. *Ann. N. Y. Acad. Sci*. 1318:71. doi: 10.1111/nyas.12464

Parasuraman, R., Sheridan, T. B., and Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Trans. Syst. Man Cybernet. A Syst. Hum.* 30, 286–297. doi: 10.1109/3468.844354

Parasuraman, R., Warm, J. S., and Dember, W. N. (1987). "Vigilance: taxonomy and utility," in *Ergonomics and Human Factors* (New York, NY: Springer), 11–32. doi: 10.1007/978-1-4612-4756-2_2

Pattyn, N., Neyt, X., Henderickx, D., and Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: boredom or cognitive fatigue? *Physiol. Behav.* 93, 369–378. doi: 10.1016/j.physbeh.2007.09.016

Quercia, A., Zappasodi, F., Committeri, G., and Ferrara, M. (2018). Local use-dependent sleep in wakefulness links performance errors to learning. *Front. Hum. Neurosci*. 12:122. doi: 10.3389/fnhum.2018.00122

Raichle, M. E., and Gusnard, D. A. (2002). Appraising the brain's energy budget. *Proc. Natl. Acad. Sci U. S. A.* 99, 10237–10239. doi: 10.1073/pnas.172399499

Raichle, M. E., and Mintun, M. A. (2006). Brain work and brain imaging. *Annu. Rev. Neurosci.* 29, 449–476. doi: 10.1146/annurev.neuro.29.051605.112819

Randall, J. G., Oswald, F. L., and Beier, M. E. (2014). Mind-wandering, cognition, and performance: a theory-driven meta-analysis of attention regulation. *Psychol. Bullet*. 140:1411. doi: 10.1037/a0037428

Ranjbar-Slamloo, Y., and Fazlali, Z. (2020). Dopamine and noradrenaline in the brain; overlapping or dissociate functions? *Front. Mol. Neurosci*. 12:334. doi: 10.3389/fnmol.2019.00334

Sara, S. J., and Bouret, S. (2012). Orienting and reorienting: the locus coeruleus mediates cognition through arousal. *Neuron* 76, 130–141. doi: 10.1016/j.neuron.2012.09.011

Sawin, D. A., and Scerbo, M. W. (1995). Effects of instruction type and boredom proneness in vigilance: implications for boredom and workload. *Hum. Factors* 37, 752–765. doi: 10.1518/001872095778995616

Scerbo, M. W. (1998). "Sources of stress and boredom in vigilance," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 42, No. 10*. (Los Angeles, CA: SAGE Publications), 764–768. doi: 10.1177/154193129804201024

Schaefer, K. E., Chen, J. Y., Szalma, J. L., and Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: implications for understanding autonomy in future systems. *Hum. Factors* 58, 377–400. doi: 10.1177/0018720816634228

Seli, P., Cheyne, J. A., Xu, M., Purdon, C., and Smilek, D. (2015). Motivation, intentionality, and mind wandering: Implications for assessments of task-unrelated thought. *J. Exp. Psychol. Learn. Mem. Cogn.* 41:1417. doi: 10.1037/xlm0000116

Seppelt, B. D., and Victor, T. W. (2016). "Potential solutions to human factors challenges in road vehicle automation," in *Road Vehicle Automation 3,* eds G. Meyer and S. Beiker (Cham: Springer), 131–148. doi: 10.1007/978-3-319-40503-2_11

Sikander, G., and Anwar, S. (2018). Driver fatigue detection systems: a review. *IEEE Trans. Intell. Transport. Syst*. 20, 2339–2352. doi: 10.1109/TITS.2018.2868499

Sirois, B., Trutschel, U., Edwards, D., Sommer, D., and Golz, M. (2009). "Predicting accident probability from frequency of microsleep events," in *World Congress on Medical Physics and Biomedical Engineering, September 7-12. 2009, Munich, Germany* (Berlin, Heidelberg: Springer), 2284–2286. doi: 10.1007/978-3-642-03882-2_607

Smallwood, J., Brown, K., Baird, B., and Schooler, J. W. (2012b). Cooperation between the default mode network and the frontal–parietal network in the production of an internal train of thought. *Brain Res*. 1428, 60–70. doi: 10.1016/j.brainres.2011.03.072

Smallwood, J., Brown, K. S., Baird, B., Mrazek, M. D., Franklin, M. S., and Schooler, J. W. (2012a). Insulation for daydreams: a role for tonic norepinephrine in the facilitation of internally guided thought. *PLoS ONE* 7:e33706. doi: 10.1371/journal.pone.0033706

Smith, G. K., Mills, C., Paxton, A., and Christoff, K. (2018). Mind-wandering rates fluctuate across the day: evidence from an experience-sampling study. *Cogn. Res. Principl. Implicat*. 3:54. doi: 10.1186/s41235-018-0141-4

Sowers, J. R., and Vlachakis, N. (1984). Circadian variation in plasma dopamine levels in man. *J. Endocrinol. Investig*. 7, 341–345. doi: 10.1007/BF03351014

Spreng, R. N. (2012). The fallacy of a "task-negative" network. *Front. Psychol*. 3:145. doi: 10.3389/fpsyg.2012.00145

Szalma, J. L., and Claypoole, V. L. (2019). "Vigilance and workload in automated systems: patterns of association, dissociation, and insensitivity," in *Human Performance in Automated and Autonomous Systems,* eds M. Mouloua, P. A. Hancock, and J. Ferraro (Boca Raton, FL: CRC Press), 85–102. doi: 10.1201/9780429458330-5

Thiffault, P., and Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. *Accid. Anal. Prev*. 35, 381–391. doi: 10.1016/S0001-4575(02)00014-3

Thomson, D. R., Besner, D., and Smilek, D. (2015). A resource-control account of sustained attention: evidence from mind-wandering and vigilance paradigms. *Perspectiv. Psychol. Sci*. 10, 82–96. doi: 10.1177/1745691614556681

Tononi, G., and Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Med. Rev*. 10, 49–62. doi: 10.1016/j.smrv.2005.05.002

Unsworth, N., and Robison, M. K. (2018). Tracking arousal state and mind wandering with pupillometry. *Cogn. Affect. Behav. Neurosci*. 18, 638–664. doi: 10.3758/s13415-018-0594-4

Van Dongen, H. P., and Dinges, D. F. (2000). Circadian rhythms in fatigue, alertness, and performance. *Principl. Practice Sleep Med*. 20, 391–399. doi: 10.1016/B0-72-160797-7/50042-2

Van Someren, E. J., Van der Werf, Y. D., Roelfsema, P. R., Mansvelder, H. D., and da Silva, F. L. (2011). "Slow brain oscillations of sleep, resting state, and vigilance," in *Progress in Brain Research, Vol. 193,* eds S. G. Waxman, D. G. Stein, D. F. Swaab, and H. L. Fields (Amsterdam: Elsevier), 3–15. doi: 10.1016/B978-0-444-53839-0.00001-6

Vyazovskiy, V. V., and Harris, K. D. (2013). Sleep and the single neuron: the role of global slow oscillations in individual cell rest. *Nat. Rev. Neurosci*. 14, 443–451. doi: 10.1038/nrn3494

Vyazovskiy, V. V., Olcese, U., Hanlon, E. C., Nir, Y., Cirelli, C., and Tononi, G. (2011). Local sleep in awake rats. *Nature* 472, 443–447. doi: 10.1038/nature10009

Wang, G., Grone, B., Colas, D., Appelbaum, L., and Mourrain, P. (2011). Synaptic plasticity in sleep: learning, homeostasis and disease. *Trends Neurosci*. 34, 452–463. doi: 10.1016/j.tins.2011.07.005

Warm, J. S., Dember, W. N., and Hancock, P. A. (1996). Vigilance and workload in automated systems. *Automat. Hum. Perform. Theor. Appl*. 1996, 183–200.

Wickens, C. (2018). Automation stages and levels, 20 years after. *J. Cogn. Eng. Decision Mak*. 12, 35–41. doi: 10.1177/1555343417727438

Wickens, C. D., Li, H., Santamaria, A., Sebok, A., and Sarter, N. B. (2010). "Stages and levels of automation: an integrated meta-analysis," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 54, No. 4* (Los Angeles, CA: Sage Publications), 389–393. doi: 10.1177/154193121005400425

Wiggins, M. W. (2011). Vigilance decrement during a simulated general aviation flight. *Appl. Cogn. Psychol.* 25, 229–235. doi: 10.1002/acp.1668

Yerkes, R. M., and Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *J. Comparat. Neurol. Psychol.* 18, 459–482. doi: 10.1002/cne.920180503

Young, R. (2012). Cognitive distraction while driving: a critical review of definitions and prevalence in crashes. *SAE Int. J. Passenger Cars-Electronic Electrical Syst.* 5, 326–342. doi: 10.4271/2012-01-0967

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroergonomics

ORIGINAL RESEARCH
published: 08 February 2022
doi: 10.3389/fnrgo.2022.787295

Let Complexity Bring Clarity: A Multidimensional Assessment of Cognitive Load Using Physiological Measures

Emma J. Nilsson[1,2]*, *Jonas Bärgman*[2], *Mikael Ljung Aust*[1], *Gerald Matthews*[3] and *Bo Svanberg*[1]

[1] Volvo Cars Safety Centre, Volvo Car Corporation, Gothenburg, Sweden, [2] Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg, Sweden, [3] Department of Psychology, George Mason University, Fairfax, VA, United States

The effects of cognitive load on driver behavior and traffic safety are unclear and in need of further investigation. Reliable measures of cognitive load for use in research and, subsequently, in the development and implementation of driver monitoring systems are therefore sought. Physiological measures are of interest since they can provide continuous recordings of driver state. Currently, however, a few issues related to their use in this context are not usually taken into consideration, despite being well-known. First, cognitive load is a multidimensional construct consisting of many mental responses (cognitive load components) to added task demand. Yet, researchers treat it as unidimensional. Second, cognitive load does not occur in isolation; rather, it is part of a complex response to task demands in a specific operational setting. Third, physiological measures typically correlate with more than one mental state, limiting the inferences that can be made from them individually. We suggest that acknowledging these issues and studying multiple mental responses using multiple physiological measures and independent variables will lead to greatly improved measurability of cognitive load. To demonstrate the potential of this approach, we used data from a driving simulator study in which a number of physiological measures (heart rate, heart rate variability, breathing rate, skin conductance, pupil diameter, eye blink rate, eye blink duration, EEG alpha power, and EEG theta power) were analyzed. Participants performed a cognitively loading n-back task at two levels of difficulty while driving through three different traffic scenarios, each repeated four times. Cognitive load components and other coinciding mental responses were assessed by considering response patterns of multiple physiological measures in relation to multiple independent variables. With this approach, the construct validity of cognitive load is improved, which is important for interpreting results accurately. Also, the use of multiple measures and independent variables makes the measurements (when analyzed jointly) more diagnostic—that is, better able to distinguish between different cognitive load components. This in turn improves the overall external validity. With more detailed, diagnostic, and valid measures of cognitive load, the effects of cognitive load on traffic safety can be better understood, and hence possibly mitigated.

Keywords: physiological measures, cognitive load, driver distraction, psychophysiology, construct validity, measurability

Edited by:
Giovanni Vecchiato,
National Research Council of Italy, Italy

Reviewed by:
Gianluca Di Flumeri,
Sapienza University of Rome, Italy
Jean-François Gagnon,
Thales Research & Technology, Canada

***Correspondence:**
Emma J. Nilsson
emma.j.nilsson@volvocars.com

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

***Received:** 30 September 2021*
***Accepted:** 07 January 2022*
***Published:** 08 February 2022*

Citation:
Nilsson EJ, Bärgman J, Ljung Aust M, Matthews G and Svanberg B (2022) Let Complexity Bring Clarity: A Multidimensional Assessment of Cognitive Load Using Physiological Measures.
Front. Neuroergon. 3:787295.
doi: 10.3389/fnrgo.2022.787295

INTRODUCTION

There are many driver states that can affect driving performance, and their contribution to risk increase will remain a central traffic safety study topic until all vehicles are fully automated. For some of these states, such as visual distraction (eyes off road) and drowsiness, the increase in risk is well-documented (Horne and Reyner, 1999; Caird et al., 2014; Victor et al., 2015). For some other states, the contribution to risk increase is less clear. One such state where the effects on driving performance are debated and further studies are needed is high cognitive load, when drivers perform non-visual, cognitively demanding activities while driving (Wijayaratna et al., 2019). (Note the difference between *cognitive demand*, which is something that is posed on the driver, and *cognitive load*, which is the resulting mental response.) It is well-established that, during increased cognitive load, response times to repeated stimuli and artificial tasks increase (Engström et al., 2010; Stojmenova and Sodnik, 2018). In addition, processing of visual information decreases (Strayer et al., 2006) and the gaze becomes more concentrated on the road ahead (Reimer et al., 2012). These findings have led to concerns about missed information and increased brake response times in critical situations (Strayer and Fisher, 2016). However, response times in unexpected critical lead-vehicle braking scenarios appear unaffected by cognitive load (Nilsson et al., 2018) and, in fact, a number of naturalistic driving studies have not found increased crash or near-crash risks for drivers talking on the phone (e.g., Fitch et al., 2013; Victor et al., 2015). It thus still remains to be sorted out when and how cognitive load poses a safety risk (see Engström et al., 2017, and Wijayaratna et al., 2019, for recent reviews and theories).

In order to study the safety impact of cognitive load, we need reliable measures that make it possible to conduct research in more naturalistic settings. The level of cognitive load can then be assessed from the measures instead of being strictly controlled by experimental manipulations. Furthermore, if future studies determine that cognitive load does indeed contribute to elevated risk in certain traffic situations, reliable measures will also be needed so that, for example, Advanced Driver Assistance Systems can detect cognitive load during driving and mitigate its effects by adapting accordingly.

A large number of studies have explored the feasibility of using physiological measures to assess cognitive load (Tao et al., 2019). Advantages of physiological measures are that they can provide continuous recordings of driver states without altering or disrupting the driving task. They can thus complement subjective and behavioral measures (which can also be very informative, but are not the focus of this article) to improve driver state assessments, or be used in situations where subjective or behavioral measures are not sensitive or appropriate (Lohani et al., 2019).

In empirical driving studies today, the level of cognitive load is usually varied systematically by having the participants perform cognitively demanding tasks (from here on referred to as cognitive tasks) while driving. It can, for example, be working memory loading tasks (Heine et al., 2017) or mental arithmetic tasks (Faure et al., 2016). The outcome (i.e., the physiological response) is then typically interpreted as reflecting the level of cognitive load. These studies might conclude, for example, that cognitive load increases the heart rate (Mehler et al., 2012) or the pupil diameter (Niezgoda et al., 2015).

This line of research has provided a great deal of knowledge regarding physiological responses to cognitive tasks. Nevertheless, there are a few well-known, yet commonly disregarded, issues that risk leading to incorrect inferences and generalizations if overlooked. In this article, we wish to bring forward these "elephants in the room," as we believe that the state of knowledge today allows greater consideration to be given to them.

Cognitive load (also often referred to as mental workload) is commonly defined as the amount of cognitive resources used to meet task demands (Engström et al., 2013; but see Van Acker et al., 2018, for a review and concept analysis). Cognitive resources enable cognitive control, which comprises neurocognitive functions resulting in effortful, conscious, and non-automatized actions (Engström et al., 2013). These functions include, for example, attention, working memory, error monitoring, and inhibitory control (Helfrich and Knight, 2016). Further, cognitive control requires cortical arousal (Grueschow et al., 2020), and can be enhanced (or degraded) by emotional responses (Critchley et al., 2013).

The first issue is thus that cognitive load consists of numerous cognitive and emotional responses that enable cognitive control during increased cognitive demand. (Cognitive and emotional responses will hereafter be jointly referred to as mental responses.) This means that cognitive load is a multidimensional construct and can take many different forms (Matthews et al., 2015). Yet, researchers almost always treat it as unidimensional when attempting to measure it.

The second issue is that task-induced cognitive load does not occur in isolation. Rather, it is part of a complex adaptation to task demands within a specific operational setting (Young et al., 2015). This issue can be best explained in two parts. First, factors other than cognitive task demand may also cause cognitive load and other mental responses, or alter the mental responses caused by the task demand (Van Acker et al., 2018). Such factors can be situation- or human-specific. Situation-specific factors characterize the context in which the task occurs and can, for example, depend on the traffic environment complexity (Törnros and Bolling, 2006; Di Flumeri et al., 2018), time pressure (Loeches De La Fuente et al., 2019), and how many times the task has been repeated (Belyusar et al., 2015). Human-specific factors include the driver's personality (Grassmann et al., 2017), experience (Paxion et al., 2014), and current mental state, such as his/her emotional state (Schoofs et al., 2008) and level of fatigue (Tanaka et al., 2009). The cognitive task and the other influencing factors together affect the driver's mental state and, consequently, his/her physiological responses and behaviors (Faure et al., 2016; Dehais et al., 2020). Second, not all mental responses to changes in cognitive demand are cognitive load components (i.e., mental responses included in the cognitive load construct). Unfortunately, it is often difficult to draw a line between those mental responses that contribute to cognitive control—and are thus to be considered cognitive load components—and those that do not. Mental fatigue is an example of a mental response to (prolonged) cognitive demand that is not part of cognitive load. Stress, on the other hand, is an example of a mental response that is difficult to categorize, since it is beneficial for cognitive control up to a certain limit, after which it has the opposite effect (Dehais et al., 2020).

Lastly, the third issue is that all physiological measures (to the best of our knowledge) correlate with multiple mental responses. These one-to-many relationships limit the inferences that can be made from the individual measures (Richter and Slade, 2017). That is, although many physiological measures correlate with cognitive load, they cannot always be considered measures of cognitive load.

Also, correlation analyses show that physiological measures that correlate with cognitive load are mostly independent of each other, implying that different physiological measures reflect different dimensions in the response to altered cognitive demand (Matthews et al., 2015). There is thus not one physiological response to cognitive load; instead, the physiological responses depend on the mental responses occurring in the specific situation at hand. Multiple physiological measures together can therefore provide us with a more comprehensive idea of the multidimensional cognitive load.

In summary, cognitive load is a complex response to cognitive demands, consisting of multiple mental responses that enable cognitive control. In empirical studies where the level of cognitive demand is altered, many different mental responses occur, depending on the cognitive task, the situation, and the individual. The different mental responses (including the different cognitive load components) have different physiological correlates, as evidenced by the lack of correlation between the physiological measures. Cognitive load can thus be neither measured, nor understood, as a unidimensional and isolated construct, and treating it as unidimensional entails a clear risk of making incorrect inferences and generalizations.

We will suggest that acknowledging these three issues is key to improving the measurability of cognitive load. Here we consider three aspects of measurability: construct validity, external validity, and diagnosticity, because of their relevance in regard to the frequently overlooked issues described above. Construct validity refers to how well a measure actually measures what it claims to measure and encompasses both the measures and the theory behind the construct (Strauss and Smith, 2009). As noted by Strauss and Smith (2009), unidimensional measures of multidimensional constructs are empirically and theoretically imprecise if the construct's components can vary independently, as is the case with cognitive load. External validity addresses the extent to which results from a study apply to other settings (Campbell and Stanley, 1963). Since traffic safety is only relevant in real life, while research on cognitive load is most often conducted in artificial environments, understanding the external validity of measures of cognitive load is highly relevant and deserves more attention than it is usually given (Jiménez-Buedo and Russo, 2021). Diagnosticity addresses a measure's ability to differentiate between different dimensions in the construct it measures (O'Donnell and Eggemeier, 1986), in this case the cognitive load components.

When designing experiments to look for physiological measures of cognitive load, one should bear in mind that the measures are typically sensitive to variations in cognitive load only within certain levels and compositions of load (de Waard, 1996), which vary for different measures and contexts. To improve the chances of finding useful measures, it is thus appropriate to look for physiological measures of cognitive load in settings where there is also an interest in understanding, and possibly mitigating, effects of cognitive load.

The cognitive control hypothesis by Engström et al. (2017) offers a plausible explanation as to how cognitive load affects driver behavior and traffic safety. It states that "*[cognitive load] selectively impairs driving subtasks that rely on cognitive control but leaves automatic performance unaffected*" (Engström et al., 2017, p. 736). Effects of cognitive load are thus relevant when exploring situations where cognitive control is required for a safe outcome; that is, in situations where drivers cannot rely solely on automated behaviors, but need to adapt their behavior using cognitive control (Engström et al., 2017). For example, cognitive load has been found to impair drivers' ability to adapt their behavior to traffic signs (Baumann et al., 2008) and downstream traffic events (Muttart et al., 2007).

On the other hand, automatized driver behaviors should not deteriorate under cognitive load, according to the cognitive control hypothesis (Engström et al., 2017). These behaviors are consistently mapped (i.e., a certain stimulus is consistently followed by the same response) and extensively practiced; one example is normal lane-keeping (Engström et al., 2017). However, if lane-keeping is made more difficult, it can be expected to require cognitive control and thus deteriorate under cognitive load (Engström et al., 2017). In line with this theory, Medeiros-Ward et al. (2014) found that when driving is made less predictable by adding wind gusts, cognitive load led to deteriorated lane-keeping. In contrast, He et al. (2014) found that cognitive load improved lane-keeping also during unpredictable wind gusts. These conflicting results call for further investigation.

As previously mentioned, a driver's mental state, including the loading on different cognitive load components, depends not only on the cognitive task but also on situation- and human-specific factors, one of which is duration. Drivers' physiological responses, behaviors, and mental states may be altered both by prolonged periods of high cognitive demand, and by underload during long-lasting tasks posing only a very low level of cognitive demand, such as simple driving (Saxby et al., 2013; Matthews et al., 2019). Desmond and Hancock (2001) named the two conditions active fatigue and passive fatigue, respectively. Although both fatigue conditions can have negative effects on task performance (Saxby et al., 2013), these effects result from different (and yet not well-understood) neurocognitive mechanisms (Berberian et al., 2017; Hu and Lodewijks, 2020) and thus require different countermeasures (Dehais et al., 2020).

Another situational factor is repetition. In experimental driving studies, tasks, and traffic scenarios are typically repeated, to increase the number of data points to improve statistical stability (Engström et al., 2010). But repeating the same cognitive tasks and traffic scenarios can lead to learning effects, which reduce the level of cognitive load and may alter its composition (Borghini et al., 2016).

Aim and Approach

The aim of this study is to demonstrate and exemplify how the measurability of cognitive load can be improved by studying multiple mental responses, using multiple physiological measures and independent variables. Changes in mental state are to be assessed based on drivers' physiological responses in relation to three independent variables, namely cognitive task demand, repetition, and traffic scenario. Five analysis questions have been defined:

Q1) How does cognitive task demand affect physiological measures?
Q2) How does repetition affect physiological measures?
Q3) Do the effects of repetition differ when the participant is just driving (baseline) compared to when also doing a cognitive task?
Q4) How do different traffic scenarios affect physiological measures?
Q5) Do the effects of traffic scenario differ when the participant is just driving (baseline) compared to when also doing a cognitive task?

Interpretations about the drivers' mental responses are to be made from answers to these questions in light of the three issues described; (1) cognitive load consists of multiple mental responses, (2) cognitive load does not occur in isolation, and (3) physiological measures correlate with multiple mental responses. The interpretations are to be based on state-of-the-art literature on physiological measures, their mental state correlations, and their neurological underpinnings. While some mental responses are clear cognitive load components, others are relevant because they affect the same physiological measures and could possibly affect the responses to the cognitive demand (e.g., Do et al., 2021).

To pursue this aim, a driving simulator study was conducted in which physiological measures were collected while cognitive task demand was manipulated with a working-memory loading n-back task at two levels of difficulty. The simulated drive consisted of a rural road with three traffic scenarios repeated four times each. When driving through these traffic scenarios, the participants were either just driving (baseline condition), or were concurrently performing the n-back task.

The following section is an overview of the physiological measures that were studied, to facilitate a nuanced discussion on the physiological responses observed in this study.

THEORY

Our bodily functions, and thus physiological responses, are controlled by the endocrine (i.e., hormonal) system and the more rapid nervous system (Tortora and Derrickson, 2007). The nervous system is divided into the central nervous system (CNS), consisting of the brain and spinal cord, and the peripheral nervous system (PNS), which connects the CNS to the rest of the body. Within the PNS, the somatic nervous system controls voluntary movements, while the autonomic nervous system (ANS) exerts involuntary, and often unconscious, control over smooth muscles, cardiac muscles, and glands (Tortora and Derrickson, 2007). The ANS is divided into the sympathetic and parasympathetic nervous systems. In general, sympathetic activation supports emergency reactions (the "fight-or-flight" response), while parasympathetic activation supports activities that occur when the body is at rest ("rest-and-digest" activities) (Tortora and Derrickson, 2007). The two systems can be co-active, reciprocally active, or independently active (Billman, 2013), and different parts of them can be activated separately (Benarroch, 2011). Several interconnected areas in the CNS integrate sensory information with emotional and cognitive processing, control the sympathetic and parasympathetic activity to maintain homeostasis, and facilitate cognitive functions and behavioral responses (Benarroch, 2011). Activity in different parts of the CNS and PNS can be observed through a variety of physiological measures.

EEG Alpha and Theta Power

Electroencephalography (EEG) is the recording of the electrical activity in the brain's outer cortex. With spectral analysis, the amount of activity (power) of different frequencies within the EEG can be studied. Increased power can be caused by repeated cycles of activation or an accumulation of transient activations (Jones, 2016). The power spectrum is typically split into the frequency bands delta (1–3 Hz), theta (4–7 Hz), alpha (8–12 Hz), beta (15–30 Hz), and gamma (30–100 Hz), although the precise frequency ranges differ between studies (Choi and Kim, 2018). The power in a certain frequency band can be studied either as absolute power or as relative power (absolute power divided by total power) (Choi and Kim, 2018). The two bands most clearly associated with cognitive load are theta and alpha.

Increases in theta power over mid-frontal cortex are frequently related to an increase in cognitive load (Cavanagh and Frank, 2014). But theta power increases also during fatigue caused by either prolonged cognitive performance (Clayton et al., 2015; Tran et al., 2020) or sleepiness (Marzano et al., 2007). In addition, surprising events give rise to transient responses within the theta frequencies (Cavanagh and Frank, 2014). The function of theta activity is not known with certainty, but it seems to reflect a need for cognitive control (Cavanagh and Frank, 2014; Cavanagh and Shackman, 2015). The need can be caused by, for example, a cognitively demanding task or a mismatch between the intended and actual level of attention—due to depleted cognitive resources, as in the case of fatigue (Clayton et al., 2015).

During task execution, alpha power generally increases in task-irrelevant sensory areas in the brain and decreases in task-relevant sensory areas (Pfurtscheller et al., 1996). During cognitive tasks, decrements in alpha power can be spread over several scalp areas (Borghini et al., 2014). Alpha power increase during mind wandering (Compton et al., 2019) and mental fatigue (Borghini et al., 2014). Previously, alpha activity was thought to represent an "idling" of the brain (Pfurtscheller et al., 1996), but theories today attribute it more functions (Halgren et al., 2019). For example, Sadaghiani and Kleinschmidt (2016) suggest that spatially widespread alpha activity contributes to tonic (i.e., slow-changing) alertness, while locally suppressed alpha activity contributes to selective attention by increasing activity and information processing in the area. The alpha band actually consists of several sub-bands, with overlapping frequencies, which respond differently to different tasks and activities (Barzegaran et al., 2017; Benwell et al., 2019).

In driving studies, cognitive task execution have led to both increased theta (frontal) and alpha (frontal, as well as more widespread) power (Sonnleitner et al., 2012; Almahasneh et al., 2014; Wang et al., 2018; Zokaei et al., 2020). During increased driving demand, alpha power has been found to decrease (Wascher et al., 2018; Abd Rahman et al., 2020), while results on theta are mixed (Wascher et al., 2018; Abd Rahman et al., 2020; Diaz-Piedra et al., 2020). Alpha and theta power both tend to increase with driver sleepiness (Simon et al., 2011; Perrier et al., 2016), although not always significantly (Ahlström et al., 2021). The effects of driving time on both bands are mixed (Perrier et al., 2016; Wascher et al., 2018; Ahlström et al., 2021). Noteworthy is that the effects of various activities or conditions can differ, depending on whether absolute or relative power measures are used (Wascher et al., 2018).

Pupil Diameter

The pupil reacts to the amount of light entering the eye by changing its size (Joshi and Gold, 2020). Additionally, cognitive and emotional states modulate the pupil size (Joshi and Gold, 2020). The pupil diameter (PD) is regulated by the sphincter muscle, which is under parasympathetic control and causes pupil constriction, and the weaker dilatory pupillary muscle, which is under sympathetic control and causes pupil dilation (Larsen and Waters, 2018). A brain area highly involved in the control the pupil size is the locus coeruleus (LC) (Joshi and Gold, 2020); the brain's primary source of the arousal-promoting neurotransmitter norepinephrine (NE) (Samuels and Szabadi, 2008). The LC-NE system plays a crucial role in cognitive processes and task performance and its activity is closely reflected by the PD (Aston-Jones and Cohen, 2005).

Pupillary responses can be studied as phasic responses and tonic levels. Phasic responses are rapid transient dilations which occur spontaneously or in response to an external stimulus (or to the lack of an expected stimulus) (Joshi and Gold, 2020). Tonic levels are studied by measuring averaged PDs, during either baseline or task conditions. A small PD indicates low vigilance or sleepiness (Zénon, 2019), while a large PD reflects high arousal (Aston-Jones and Cohen, 2005) or high levels of cognitive activity (Zénon, 2019). During task execution, a large PD indicates greater effort and often correlates with good performance (van der Wel and van Steenbergen, 2018).

Numerous driving simulator studies have shown increased PD during cognitively (Hammel et al., 2002; Niezgoda et al., 2015; Cegovnik et al., 2018; He et al., 2019; Peruzzini et al., 2019) and visually (Benedetto et al., 2011) demanding secondary tasks, psychological stress (Pedrotti et al., 2014), and time pressure (Rendon-Velez et al., 2016), as well as during increased driving demand (Peruzzini et al., 2019; Xie et al., 2020). As task demands increase, the PD typically shows a stepwise increase before it plateaus or decreases again at high load levels when performance can no longer be maintained (van der Wel and van Steenbergen, 2018). The plateau and decrease are likely due to a decrement in motivation and effort (van der Wel and van Steenbergen, 2018).

Few studies have explored the effects of secondary tasks on pupil diameter in real driving. Because the pupil is very sensitive to lighting variations, task effects risk being masked in real-life environments with fluctuating light levels. Nonetheless, Nunes and Recarte (2002) and Recarte and Nunes (2000) found that the PD increased during the execution of cognitive tasks on real roads, except during simple conversation tasks (Nunes and Recarte, 2002). Further, Ahlström et al. (2021) found a decrease in PD with increased distance driven by sleep-deprived drivers at nighttime.

Eye Blink Rate and Duration

Eye blinks are essential for lubricating the eyes, but characteristics such as their frequency and timing depend on cognitive and emotional factors as well (Cruz et al., 2011). The eye blink rate (EBR) is positively related to the level of the neurotransmitter dopamine in the brain (Eckstein et al., 2017), although the precise relationship is unknown (Jongkees and Colzato, 2016; Sescousse et al., 2018). Dopamine affects several brain functions, including cognitive control, motivation, and learning (Jongkees and Colzato, 2016). Levels of dopamine that are too low or too high, reflected in low or high EBR, typically mean worse performance (Jongkees and Colzato, 2016; Eckstein et al., 2017) due to depressed prefrontal cortex activation (Dehais et al., 2020).

Brain activity studies have suggested that spontaneous eye blinks provide brief moments of attentional disengagement from an external stimulus in favor of internal processing (Nakano et al., 2013). Blinks occur less frequently during visually demanding tasks, probably to reduce the risk of missing relevant information (Recarte et al., 2008). This reduction in frequency has been demonstrated in laboratory studies (Recarte et al., 2008; Cardona et al., 2011) and in driving studies investigating increased driving demand (Wiberg et al., 2015; Faure et al., 2016; Lobjois et al., 2021). In driving studies applying visually demanding secondary tasks, the effect has not reached significance (Liang and Lee, 2010; Benedetto et al., 2011). Unfortunately, because large saccades (quick movements of both eyes) are often accompanied by blinks (Fogarty and Stern, 1989), comparing EBR between different traffic environments or tasks with different glance behaviors can be problematic (Cardona and Quevedo, 2014).

During cognitive tasks, the EBR increases both in laboratory (Recarte et al., 2008; Magliacano et al., 2020) and driving studies (Nunes and Recarte, 2002; Liang and Lee, 2010; Niezgoda et al., 2015; Faure et al., 2016; Cegovnik et al., 2018; He et al., 2019; Chihara et al., 2020). Although these results are highly consistent, EBR differences between levels of cognitive load are typically small and rarely significant.

The effects of increased visual and cognitive demands on eye blink duration (EBD) are less explored and less consistent. Simulator studies have not found significant effects of either traffic complexity (Faure et al., 2016) or cognitively (Faure et al., 2016) or visually (Benedetto et al., 2011) demanding secondary tasks on EBD. However, studies in real traffic have demonstrated shorter EBDs in drivers compared to their passengers (Takeda et al., 2016), as well as during driving in more demanding traffic situations (Wiberg et al., 2015).

During increased driver sleepiness, EBD increase consistently (Ahlström et al., 2018; Cori et al., 2019). The studies are fewer and results ambiguous regarding the effects of driver sleepiness on EBR (Cori et al., 2019). On the contrary, during increased fatigue due to prolonged task execution, the pattern is reversed: studies consistently show increased EBR while the effects on EBD are mixed (Bafna and Hansen, 2021).

Heart Rate and Heart Rate Variability

The heart is regulated through both the sympathetic and the parasympathetic nervous systems. Sympathetic activity causes the heart to beat faster and stronger, while parasympathetic activity decelerates the heart rate (HR) (Tortora and Derrickson, 2007). The systems can be activated individually or simultaneously and in the same or opposite directions, but parasympathetic activity is faster and stronger than sympathetic activity (Billman, 2013). A healthy heart has a constantly fluctuating HR (Park and Thayer, 2014), measured as heart rate variability (HRV). The respiratory cycle has a major influence on HRV, as inhalations accelerate the heart and exhalations slow it down (Quintana and Heathers, 2014).

Cardiac activity is most often described by HR and HRV. HRV is an umbrella term for different measures of the fluctuations in the time intervals between adjacent heart beats (for an overview, see Shaffer and Ginsberg, 2017). One common HRV measure (used in this study) is the root mean square of successive differences (RMSSD), which is supposed to reflect parasympathetic activity without much respiratory influence (Laborde et al., 2017).

Cognitive tasks lead to increased HR and decreased HRV in laboratory environments (Hughes et al., 2019), as well as in driving studies in simulators (Belyusar et al., 2015; Hidalgo-Muñoz et al., 2019; Tejero and Roca, 2021) and on real roads (Reimer and Mehler, 2011; Mehler et al., 2012). The effects of driving demand on HR and HRV are however varying. For example, simulator studies by Foy and Chapman (2018) and Stuiver et al. (2014) didn't find any effect of varying driving demands on HR, while Wiberg et al. (2015) did find such an effect in real city driving—but the result was less consistent in highway driving. Further, Dussault et al. (2004) found increased HR in pilots during active flight segments compared to in-flight rest segments, but only during actual (not simulated) flights (Dussault et al., 2005). In a driving simulator study by Beggiato et al. (2019), participants' HR typically decreased when approaching traffic scenarios designed to evoke unease. This could be a sign of attentional focusing and preparation for action as HR decelerations are known to occur in aiming sports before an athlete throws a dart or makes a golf putt, for example (Cooke, 2013).

The effects of prolonged task execution and increased mental fatigue on HR and HRV are inconclusive; in fact, both increased (Matuz et al., 2021) and decreased (Li et al., 2002; Mizuno et al., 2011) parasympathetic activity has been suggested. In driving studies, sleepiness due to sleep deprivation causes HR to decrease and HRV to increase on a group level, but individual variation is large (Buendia et al., 2019; Persson et al., 2020; Ahlström et al., 2021).

In general, emotions characterized by passivity, such as sadness, contentment, and suspense, cause a decrease in HR, whereas the opposite is true for emotions characterized by active coping responses, such as anger, embarrassment, and fear (Kreibig, 2010). As an example, HR increases during emotional stress caused by having one's performance judged (Kelsey et al., 2004). The effects of emotions on HRV are less consistent (Kreibig, 2010). Responses to novel stimuli cause a temporary HR deceleration (Bradley, 2009), due to co-activation of the slower sympathetic and faster parasympathetic systems (Silvani et al., 2016).

Breathing Rate

Breathing, which is under both voluntary and involuntary control (Homma and Masaoka, 2008) both affects, and is affected by, emotions and cognition (Homma and Masaoka, 2008; Del Negro et al., 2018). Roughly every fifth minute, rhythmic breathing is interrupted by a sigh (Del Negro et al., 2018). Sighs open up collapsed alveoli (Del Negro et al., 2018) and reset the breathing rhythm (Vlemincx et al., 2012). Sighs also occur in response to emotions such as grief and happiness (Del Negro et al., 2018), and cause emotional relief (Vlemincx et al., 2013).

The most frequently studied breathing measure in studies of cognitive load is breathing rate (BrR), which consistently increases during cognitive task execution (Grassmann et al., 2016). In single task studies, BrR has also been successful in discriminating between different levels of cognitive load (Backs and Seljos, 1994; Brouwer et al., 2014; Hogervorst et al., 2014; Hidalgo-Muñoz et al., 2019), but this load level sensitivity seems to disappear in driving studies (Mehler et al., 2009; He et al., 2019; Hidalgo-Muñoz et al., 2019). The effects of traffic complexity on BrR appear inconsistent: Wiberg et al. (2015) found BrR to increase during increased traffic complexity in real driving, while Foy and Chapman (2018) found no such effect in a simulator study.

When drivers are sleepy, BrR has been shown to decrease (Kiashari et al., 2020) and become less regular (Rodrígue-Ibáñez et al., 2011). The few studies that have looked at the effects of prolonged execution of cognitive tasks on BrR show inconsistent results (see Grassmann et al., 2016, for a review). BrR also increases during time pressure (Rendon-Velez et al., 2016) and as a result of emotions such as anxiety (Homma and Masaoka, 2008), fear (Stephens et al., 2010), and amusement (Stephens et al., 2010), while it decreases as a result of calm and positive emotions (Balters and Steinert, 2015).

Skin Conductance

Electrodermal activity is the change in the electrical properties of the skin, typically measured as skin conductance (SC). As sweat ducts fill with sweat, the resistance of the outer layer of the skin decreases and the conductance increases (Dawson et al., 2016). The sweat glands on the palms and soles are densely distributed and primarily respond to emotional arousal in what is known as psychological or emotional sweating (Baker, 2019). The function of emotional sweating is likely to improve grasping as part of the body's preparation to act or flee (Dawson et al., 2016). Sweating is regulated by the sympathetic nervous system alone, making SC a popular measure of general arousal (Posada-Quintero and Chon, 2020).

The SC is most often quantified as tonic changes of skin conductance level (SCL) and phasic sweat bursts called skin conductance responses (SCRs). SCRs occur as part of the orienting response when attention is directed toward a novel, significant stimulus (Bradley, 2009), and also follow deep breaths and body movement (Dawson et al., 2016). In addition, they occur spontaneously approximately one to three times per minute during rest (Dawson et al., 2016).

In principle, the anticipation and performance of practically any task invoke increased SC (in both SCL and SCRs) (Dawson et al., 2016). Cognitive tasks cause increased SC in laboratory settings (Brouwer et al., 2014; Visnovcova et al., 2016) as well as in driving studies in simulators (He et al., 2019) and real cars (Reimer and Mehler, 2011; Mehler et al., 2012, 2016). The effect of performing a task (compared to a baseline condition) is often greater than the differences between task load levels (Reimer and Mehler, 2011; Mehler et al., 2016), and recovery to baseline levels is rather slow (Mehler et al., 2012; Visnovcova et al., 2016). Studies of driving demand demonstrate that increased traffic complexity leads to increased SC in real (Wiberg et al., 2015) and simulated driving (Foy and Chapman, 2018).

Few studies have been conducted on the effects of sleepiness on SC. Although a decrease in SCR frequency (Michael et al., 2012) and SCL (Miró et al., 2002) has been demonstrated due to sleep deprivation, the effects are rather small and inconsistent—and accompanied by stronger circadian oscillations. As for the effects of emotions, SC typically increases in response to those emotions high in arousal (Kreibig, 2010; Gomez et al., 2016). It has been suggested that the increase in SC reflects motor preparation, as many emotions call for action (Kreibig, 2010). This interpretation explains why emotions related to passivity, such as contentment, relief, and sadness, show decreased SC (Kreibig, 2010).

METHOD

The study consisted of two similar test series, Test Series 1 and Test Series 2. Differences consisted of the cognitive tasks employed, and in the design of one of the traffic scenarios (see descriptions in Sections Cognitive Task and Driving Scenarios). Otherwise, the test series were the same. Data were collected at the Swedish National Road and Transport Research Institute (VTI) in Linköping, Sweden. The experiment was approved by the regional ethics vetting board (Regionala etikprövningsnämnden) in Linköping.

Participants

Participants were recruited from a random selection of the vehicle register over people living in the Linköping area. A total of 70 males participated in the study, 36 in Test Series 1 and 34 in Test Series 2. They ranged in age from 35 to 51 years ($M = 43$, $SD = 4$), drove between 50 and 1,200 km/week ($M = 309$, $SD = 205$), and had held a driver's license for between 8 and 32 years ($M = 23$, $SD = 5$). Additional requirements for participating in the study were to: have normal hearing; have a BMI < 30; not rate oneself as extreme in extraversion or introversion, stress-sensitivity, and morning- or evening-type; not have bad health or use medication regularly; not have sleep disorders; and be able to abstain from nicotine for 3 h without withdrawal symptoms. The requirements were there to create a fairly homogenous group of participants to reduce the variance in both mental and physiological responses to the experimental manipulations.

All participants were paid 1500 SEK for their participation.

Equipment

The experiments were carried out in an advanced moving-base driving simulator. The car body consisted of the front part of a SAAB 9-3 with automatic transmission mounted on a cradle which allowed movement with four degrees of freedom. The field of vision was 120°, and three LCD displays were used to simulate rear-view mirrors. A sound system simulated sounds from the tires and engine. The test leader could communicate with the participants through speakers, which were also used for the cognitive tasks.

Electrooculography (EOG), electroencephalography (EEG), electrocardiography (ECG), electromyography (EMG), skin conductance (SC), and respiratory inductance plethysmography (RIP) signals were recorded using a multi-channel amplifier (g.HIamp, g.tec Medical Engineering GmbH, Austria). Thirty-two EEG channels (Fp1, FpZ, Fp2, F7, F3, FZ, F4, F8, FC5, FC1, FC2, FC6, T7, C3, CZ, C4, T8, CP5, CP1, CP2, CP6, P7, P3, PZ, P4, P8, POZ, O1, OZ, O2, A1, A2) and four EOG channels were recorded using active electrodes on a cap (g.tec g.GAMMAcap), referenced to the right earlobe (A2), and with a ground electrode at AFZ. The EEG electrodes were positioned according to the 10–20 system. Two EOG electrodes were placed horizontally outside the outer canthus of each eye, and two were placed vertically across the left eye. The ECG was recorded with electrodes placed on the right collarbone and a lower left rib. The SC was recorded from the distal phalanges at the index and middle fingers at the left hand, and the RIP with an elastic strap placed around the participant's chest, just below the armpits. The EMG was recorded with electrodes placed on the trapezius (shoulder) and masseter (jaw) muscles. The EMG data was collected for the purpose of EEG artifact handling, but because it was not found to be useful for that purpose, EMG was not included in the analysis and will not be described further. All physiological signals were recorded with a sampling rate of 256 Hz. The EEG, EOG, and ECG signals were band-pass filtered between 0.5 and 60 Hz using an 8th order Butterworth filter and notch filtered between 48 and 52 Hz using a 4th order Butterworth filter. The SC and RIP signals were band-pass filtered between 0 and 30 Hz using an 8th order Butterworth filter.

The pupil diameter was measured with a Smart Eye four-camera system in Test Series 1, and with eye tracker glasses from SensoMotoric Instuments (SMI) in Test Series 2.

Cognitive Task

The cognitive task was an auditory, non-verbal version of the n-back task (see Mehler et al., 2011, for a similar verbal version).

It is well-established that n-back tasks cause increased levels of cognitive load (Jaeggi et al., 2010). A number between zero and nine was orally presented to the participants every other second. If the number just presented was the same as the number presented n numbers ago, it was considered a target number. All number series were unique, consisting of 30 numbers, with six target numbers. The participants were instructed to press a button mounted on their right index finger against the steering wheel as soon as they detected a target number. In Test Series 1, that task was only presented at the 1-back ($n = 1$) level, while in Test Series 2, the task was presented at both the 1-back and 2-back ($n = 2$) levels. Right before the task began, the participants were informed through the speakers that the task would begin, and which level it would be.

Driving Scenarios

The simulated driving environment consisted of a two-lane rural road with a speed limit of 80 km/h. There was occasional traffic, both oncoming and overtaking.

Measurements were collected during three traffic scenarios (Hidden Exit, Intersection, and Wind), each repeated four times during the drive: see **Figure 1**. In the Hidden Exit scenario, a warning sign for a hidden exit was placed before a sharp right curve with a high hedge on the right (inner) side. After the curve was the exit on the right side. There was no other traffic in the scenario. In the Intersection scenario, the participants approached and drove through a four-way intersection, in which they had the right of way. Another car approached the intersection from the right, becoming visible as it drove past a house when the participants were 180 m from the intersection, and came to a stop at the intersection 2 s before the participants reached the intersection. A bus in the oncoming lane passed the participant's car 70 m before the intersection. In the Wind scenario, the otherwise present forest surroundings opened up into a field with very limited road curvature. While driving through the field, the participants were occasionally exposed to crosswinds from the right. Wind speed was determined by three overlaid sinusoidal winds with different frequencies, resulting in a, for the participants, unpredictable crosswind.

During the measurement scenarios, the participants were either engaged in the 1-back task or the 2-back task or they were just driving (the baseline condition). In the Hidden Exit and Intersection scenarios, the n-back task started ~45 s before the participants reached the hidden exit or intersection. In the Wind scenario, the crosswinds started 30 s before the task onset and blew for 1 min and 40 s.

The two test series differed somewhat in their design. In Test Series 1, the crosswinds were always active in the Wind scenario. For two repetitions of each scenario the participants were engaged in the 1-back task, and for the other two repetitions, there was no task besides driving (baseline). In Test Series 2, the crosswinds were only active in two of the four Wind scenario repetitions. In each crosswind condition (Wind On and Wind Off), the participants performed the 2-back task once and the baseline condition once. In the Hidden Exit and Intersection scenarios, the participants were engaged in the 1-back task once, the 2-back task twice, and baseline once. The order of the tasks was counterbalanced across participants in both series.

In addition to the measurement scenarios, there were some other scenarios that only differed from the measurement scenarios in terms of traffic, for the sake of variation. There were two more hidden exits with a car standing still at the exit with indicators on, two four-way intersections with a car approaching from the left, and two four-way intersections with no other traffic.

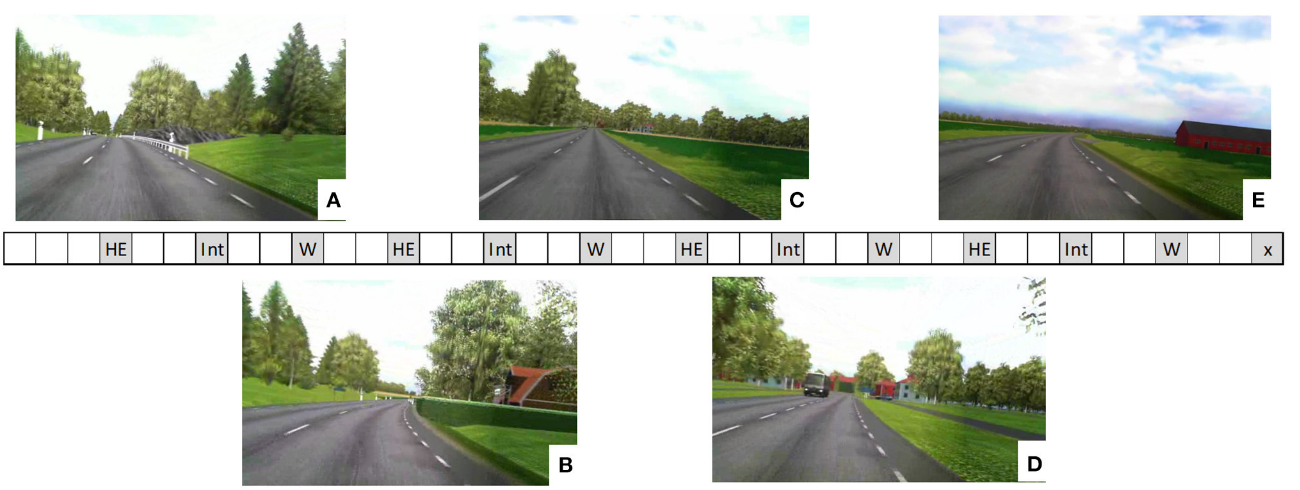


FIGURE 1 | Illustration of the experimental drive. The rectangle in the middle of the figure represents the sequence of events. Each square is ~1 min of driving (depending on vehicle speed). Measurement scenarios are colored in gray. HE, Hidden Exit; Int, Intersection; W, Wind. The final scenario, x, was an unexpected lead vehicle braking scenario which is not included in this study (but see Nilsson et al., 2018). The five images show scenery examples from one of the positions where the hidden exit warning sign became visible for the driver **(A)**; the hidden exit became visible **(B)**; 500 m before the 4-way intersection **(C)**; 80 m before the 4-way intersection **(D)**; and the open field in the Wind scenario **(E)**. Images are blurry due to poor camera resolution, but the participants experienced them in high resolution in the driving simulator.

Procedure

Prior to their participation in the study, participants were sent a background questionnaire and a written description of the study. They were asked to abstain from alcohol for 72 h and nicotine and caffein for 1 h before the experiment. On arrival at the laboratory, they handed in the questionnaire, and the study was once again explained. After the participants gave their written consent to participating in the study, the physiological monitoring equipment was attached. They were then taken to the simulator, the equipment was connected to the measurement devices, and the eye tracker was calibrated.

The participants practiced the 1-back and 2-back (Test Series 2 only) tasks until they felt comfortable performing them. After being informed about the simulator, the participants drove for ∼10 min to practice driving. In Test Series 1, the practice drive included two 1-back tasks. In Test Series 2, it included one 1-back and two 2-back tasks. In both series, the practice drive also included one hidden exit scenario with a car standing still at the exit with its indicator on, and one four-way intersection scenario identical to the Intersection scenario.

During the practice session, the participants were allowed to speak to the test leader. Then followed the actual experiment, which lasted ∼40 min, during which the participants were asked not to talk to the test leader unless it was urgent. After the drive, participants filled out a questionnaire about their test experience. The overall time, from when participants arrived until they left, was ∼3.5 h.

Physiological Measures

In each measurement scenario, each physiological measure was derived as one averaged value and one continuous vector. The averaged values were computed over 50 s, starting 10 s after task onset. The first 10 s were excluded to reduce the effects of the surprise reaction at task onset. The continuous vectors were initially derived with constant time steps (described for each measure below). The vectors were then transformed to having constant distance steps instead, so that they could be visualized in relation to the traffic environment. The data processing was done in MATLAB R2015b and MATLAB R2019b.

All signals and derived measures in each analyzed segment were visually inspected to ensure adequate data quality before being included in the analysis (a slightly different procedure for the EEG measures is described below). For each measure and scenario, participants were included in the analysis only if they had a complete dataset of all four repetitions.

For EEG data processing, the MATLAB toolbox EEGLAB vs 2020.0 (Delorme and Makeig, 2004) was used. Ninety-second EEG segments, starting 10 s before task onset and ending 20 s after task end (or corresponding segments in the baseline conditions), were extracted. For each segment, all EEG channels were visually inspected. Channels with poor signal quality (either high levels of high frequency noise throughout the recording, or that contained large or frequent signal deviations) were removed. The average number of remaining channels was 27.7 (std 1.5). Also, epochs that contained large movement or muscle artifacts were removed (Tatum, 2014). The remaining channels were then re-referenced to linked ears. To suppress remaining artifacts, which were primarily caused by eye blinks and eye movements, independent component analysis (ICA) was performed on the data, using Infomax ICA (runica, Makeig et al., 1996). The resultant independent components (ICs) were classified into seven categories, including "brain activity" and "eye activity," using the default classifier in ICLabel (Pion-Tonachini et al., 2019). ICs that were classified as having <20% brain activity or >70% eye activity were removed. The average number of removed ICs was 12.4 (std 3.4). After testing different thresholds on a randomly selected subset of EEG segments, we chose values that led to the exclusion of evident artifacts while retaining as much data as possible. The remaining ICs were subsequently transformed back to the EEG channels. For the averaged values, power spectra were then calculated using Welch's power spectral density estimate with a 2-s window, 50% overlap, and windowing using a Hamming window. The average power was derived for channels F3, FZ, and F4 in the 4–7.5 Hz theta frequency range and for channels P3, PZ, and P4 in the 8–13 Hz alpha frequency range. These averaged power values were then divided by the sum of the total power in the 4–25 Hz frequency range in the same channels, resulting in a relative frontal-midline theta power (Theta) value, and a relative parietal-midline alpha power (Alpha) value. Because artifacts were handled for each segment separately, what ICs were derived and which ones were removed differed between segments. This caused some added variation in absolute power in the processed channels between the segments. Relative, rather than absolute, measures were therefore used as they were less affected by these segment variations. Continuous Theta and Alpha were derived using the same method, but for one 2-s segment at a time, moving in 1/8-s steps, to make continuous vectors. The Alpha vectors were transformed with the natural logarithm to achieve an approximately normal distribution. Alpha and Theta were only derived for the segments in Test Series 2 due to lack of time.

R-peaks were detected in the ECG signals using the *qrsdetect* function in the Biosig toolbox (Vidaurre et al., 2011), and the R-R-intervals (RRIs) were derived as the time between adjacent heart beats. To remove abnormal or artifactual heart beats, RRIs that differed more than 30% from the surrounding six RRIs were removed (Karlsson et al., 2012). The RRIs were then converted to heart rate (HR; beats/min). The continuous HR vectors were derived by linearly interpolating the discrete HR values. Finally, the continuous HR vectors were normalized by subtracting the entire drive's median HR value to reduce between-subject variance in the continuous plots.

The HRV was computed as the root mean square of the successive differences in RRIs after artifact removal (RMSSD) for each 50-s analysis segment (Shaffer and Ginsberg, 2017). The RMSSD values were then log transformed using the natural logarithm to make the distribution more normal (Laborde et al., 2017). No continuous HRV vector was made.

Breaths were detected in the RIP signal using an in-house algorithm based on local peaks and thresholds, and the mean breathing rate (BrR; breaths/min) was derived by counting the number of breaths in the segment. For the continuous BrR vectors, the time between adjacent breaths was derived and converted to breaths per minute. The discrete BrR values were interpolated with next-neighbor interpolation and the vector was normalized by subtracting the entire recording's median BrR value.

In driving studies, SC is commonly studied by averaging of the signal (sometimes after artifact removal or normalization) over relevant segments in time (Mehler et al., 2010; e.g., Loeches De La Fuente et al., 2019). Here, the SC data required some additional signal processing to achieve normal distributions and remove effects of signal drift (which was evident from visual inspection of the signals, as advised by Braithwaite et al., 2015). The following processing steps were conducted. The SC signals were first smoothed with a function, *wsmooth*, based on the Whittaker's smoother (Eilers, 2003), and then filtered using a 2nd order Butterworth lowpass filter with cutoff frequency 0.001 Hz. By subtracting the filtered signal from the SC vector, an SCR vector was derived. The SC vector was then divided by the SCR vector's 99th percentile value (representing that participant's SCR amplitude). The 99th percentile and not the maximum value was used to avoid influence of any rare extreme values or artifacts. This individual response amplitude normalization was necessary to achieve normal distributions. Finally, to compensate for the drift, the vector's average value in the interval 70–10 s before task onset was subtracted from the analysis segment (principle described in Geršak, 2020).

Eye blinks were detected in the vertical EOG signal with an algorithm based on derivatives and thresholds (Jammes et al., 2008). The mean eye blink rate (EBR; blinks/min) was derived by counting the number of eye blinks in the segment. For the continuous EBR vectors, the time between adjacent eye blinks was derived and converted to blinks per minute. The entire recording's median EBR was then subtracted from all EBR data points to reduce interindividual differences. Next, a constant was added to all the data points to make them ≥ 1, after which they were transformed using the natural logarithm to make them more normally distributed (as suggested by Cruz et al., 2011). Finally, the discrete EBR values were interpolated using next-neighbor interpolation.

Eye blink duration (EBD; ms) was defined as the time between the eye blink's half rise amplitude and half fall amplitude to reduce the problem with otherwise hard-to-define start and end times (as e.g., Ahlström et al., 2018). Eye closures with a duration >500 ms (considered non-blink closures in International Organization of Standardization, 2014) were excluded from the analysis to avoid extreme outliers. The continuous EBD vectors were derived with the same procedure as the continuous EBR vectors.

The pupil diameter (PD; mm) was obtained from the eye trackers. Sudden drops in the PD vector were removed through linear interpolation, to reduce the effects of eye blinks and other tracking issues (Klingner, 2010). In Test Series 2, one PD vector was obtained for each eye and the one with the best signal quality, assessed through visual inspection, was used in the analysis (unless they were both excluded due to poor signal quality). In Test Series 1, only one PD vector was obtained for each subject. The absolute PD values differed between the series, due to the different eye trackers. Both the averaged PD values and the continuous PD vectors were thus normalized by subtracting the entire recording's median PD value. After this normalization, there were no longer any statistically significant differences in absolute PD values between the two series for the same scenarios and task conditions.

Statistical Analysis

Each physiological measure was analyzed using a mixed model ANOVA with task (baseline, 1-back, and 2-back) and traffic scenario (Hidden Exit, Intersection, and Wind) as categorical fixed-effect variables, scenario repetition (1–4) as a quantitative fixed-effect variable, and test participant as a categorical random-effect variable. Two-way interactions between task and traffic scenario and between task and repetition were included in the model. The significance level was set to 0.05 and Bonferroni correction was used to compensate for the multiple tests. The normality assumption of each ANOVA was confirmed by controlling that its residuals followed an approximately normal distribution (see **Supplementary Material**; **Appendix 1**).

In addition, using data from the Wind scenario in Test Series 2, the effects of the crosswinds were tested separately for the task conditions baseline and 2-back (recall that there was no 1-back condition in the Wind scenario in Test Series 2). A mixed model ANOVA was used, with crosswind (Wind on, Wind off) as a categorical fixed-effect variable and test participant as a categorical random-effect variable.

Effects of traffic scenarios were further explored with continuous plots of mean values and their corresponding 95% confidence intervals (CIs), similar to Beggiato et al. (2019). The distributions of the data samples were approximately normally distributed around the means. Note, however, that because the data samples in a continuous plot are not independent from each other, non-overlapping confidence intervals does not necessarily imply a statistically significant difference between two points (Cumming and Finch, 2005). Therefore, paired *t*-tests were made between two points in time (Wind scenario), or position (Hidden Exit and Intersection scenario), for each task condition. The points were chosen so that the level of demand from the traffic scenario was assumed to differ between them, and so that most of the related responses that were visible in the plots took place between them. In the Wind scenario, two tests were made between the point in time where the two greatest wind bursts occurred, compared to 7 s earlier, where the wind was low. In the Hidden Exit scenario, one test was made between the position where the warning sign first became visible to the participant, and the position where the hidden exit first became visible. In the Intersection scenario, one test was made between the position 80 m before the intersection, where the approaching car had slowed down and was approximately one car length from the stopping point, and the position 500 m before the intersection. The position 500 m before the intersection was chosen because it was not clear at what position the participants recognized the scenario, and so a rather large distance to the more demanding part of the scenario was chosen. Examples of what these analysis positions could look like can be seen in **Figure 1**.

At these points, the average value for each of the physiological measures' continuous vectors in a 2.1 s (Wind scenario), or 42 m (Hidden Exit and Intersection scenario), interval, centered around the analysis point, was derived for each participant and repetition. These averaged values were then used in the paired *t*-tests. Because multiple *t*-tests were made, and the points for testing were selected after the data had been visualized, results need to be interpreted with caution as the risk of type 1 errors is inflated (Forstmeier et al., 2017). Because we want to avoid inflating the risk of type 2 errors and missing actual effects, correction for multiple testing has not been made (Forstmeier et al., 2017). Instead, consistency and effect sizes of visualized and statistically tested effects are considered in the result interpretations.

The mixed model ANOVAs were performed using SAS Enterprise Guide 8.2 and continuous plots and *t*-test with MATLAB R2019b.

RESULTS

Four of the 70 participants aborted the experiment due to simulator sickness and 3 were excluded from the analysis due to data loss in the logging system. A total of 63 participants were hence included in the analysis. Of these, 9 lacked a complete PD dataset due to logging issues. For one participant in Test Series 2, the n-back task did not start as intended in one Wind scenario, so this participant has three occasions of baseline and only one occasion of 2-back in the four Wind scenarios.

Crosswind

The mixed model ANOVAs revealed no significant effect of crosswinds in any of the physiological measures, either in the baseline condition or in the 2-back condition: see **Table 1**. The two crosswind conditions (Wind On and Wind Off) were therefore merged in the remaining analyses.

Q1) How Does Cognitive Task Demand Affect Physiological Measures?

The mean values, standard deviations, and the number of samples included are presented for each measure in each task condition (all repetitions and scenarios are merged) in **Table 2**.

Detailed results from the mixed model ANOVAs of the effects of task, repetition, and scenario are presented in **Table 3**. The task had a significant effect on HR, RMSSD, BrR, SC, PD, and EBR in the form of a stepwise increase (or decrease) with increasing level of cognitive demand in all measures. Only in EBR was the difference between 1-back and 2-back tasks not significant. There was no significant effect of task in EBD, Alpha, or Theta.

Q2) How Does Repetition Affect Physiological Measures?

Repetition had a significant effect on HR, BrR, PD, EBR, EBD, and Alpha, but not on RMSSD, SC, and Theta (see **Table 3**).

TABLE 1 | Effects of crosswinds on each physiological measure and task condition.

	Baseline			2-back		
	Wind On m (sd)	Wind Off m (sd)	Main effect	Wind On m (sd)	Wind Off m (sd)	Main effect
HR (beats/min)	65.18 (9.34)	65.69 (9.47)	$F_{(1,30)} = 1.47$, $p = 0.23$	68.79 (9.61)	68.55 (10.01)	$F_{(1,28)} = 0.68$, $p = 0.42$
RMSSD (-)	3.57 (0.51)	3.53 (0.48)	$F_{(1,30)} = 0.40$, $p = 0.53$	3.18 (0.48)	3.19 (0.46)	$F_{(1,28)} = 0.03$, $p = 0.87$
BrR (breaths/min)	15.41 (3.35)	15.60 (3.83)	$F_{(1,19)} = 0.15$, $p = 0.70$	19.14 (3.45)	19.27 (4.36)	$F_{(1,17)} = 0.02$, $p = 0.90$
SC (–)	−0.119 (0.143)	−0.011 (0.221)	$F_{(1,22)} = 3.81$, $p = 0.06$	0.105 (0.314)	0.089 (0.355)	$F_{(1,20)} = 0.01$, $p = 0.91$
PD (mm)	−0.343 (0.123)	−0.410 (0.171)	$F_{(1,18)} = 3.26$, $p = 0.09$	0.090 (0.193)	0.036 (0.245)	$F_{(1,18)} = 1.84$, $p = 0.19$
EBR (blinks/min)	32.66 (13.97)	34.85 (13.02)	$F_{(1,22)} = 2.97$, $p = 0.10$	36.26 (19.01)	38.19 (19.35)	$F_{(1,22)} = 1.44$, $p = 0.24$
EBD (ms)	123.9 (23.4)	130.6 (25.2)	$F_{(1,22)} = 3.50$, $p = 0.07$	123.7 (33.3)	125.4 (35.7)	$F_{(1,22)} = 0.15$, $p = 0.70$
Alpha (–)	0.0325 (0.0096)	0.0321 (0.0087)	$F_{(1,23)} = 0.07$, $p = 0.79$	0.0324 (0.0117)	0.0320 (0.0095)	$F_{(1,21)} = 0.22$, $p = 0.65$
Theta (–)	0.0510 (0.0121)	0.0491 (0.0116)	$F_{(1,23)} = 1.30$, $p = 0.26$	0.0559 (0.0130)	0.0554 (0.0129)	$F_{(1,21)} = 0.07$, $p = 0.79$

No correction has been done to compensate for multiple tests to reduce the risk of type 2 errors. HR, heart rate; RMSSD, root mean square of successive differences between heart beats; BrR, breathing rate; SC, skin conductance; PD, pupil diameter; EBR, eye blink rate; EBD, eye blink duration; Alpha, relative EEG alpha power; Theta, relative EEG theta power; m, mean; sd, standard deviation.

TABLE 2 | Measure statistics.

	Baseline	1-back	2-back
HR (beats/min)	m = 64.87 sd = 9.37 *n* = 312	m = 67.30 sd = 10.22 *n* = 251	m = 69.04 sd = 10.31 *n* = 177
RMSSD (–)	m = 3.56 sd = 0.48 *n* = 312	m = 3.39 sd = 0.51 *n* = 251	m = 3.20 sd = 0.49 *n* = 177
BrR (breaths/min)	m = 16.96 sd = 3.71 *n* = 219	m = 19.06 sd = 3.06 *n* = 180	m = 19.16 sd = 3.42 *n* = 117
SC (–)	m = −0.030 sd = 0.241 *n* = 270	m = 0.068 sd = 0.249 *n* = 225	m = 0.176 sd = 0.302 *n* = 133
PD (mm)	m = −0.141 sd = 0.246 *n* = 188	m = 0.147 sd = 0.272 *n* = 150	m = 0.334 sd = 0.293 *n* = 114
EBR (blinks/min)	m = 29.21 sd = 12.41 *n* = 244	m = 31.97 sd = 14.27 *n* = 198	m = 35.35 sd = 18.54 *n* = 142
EBD (ms)	m = 122.0 sd = 26.5 *n* = 244	m = 123.9 sd = 32.9 *n* = 198	m = 117.2 sd = 30.0 *n* = 141
Alpha (–)	m = 0.0317 sd = 0.0091 *n* = 92	m = 0.0302 sd = 0.0082 *n* = 45	m = 0.0308 sd = 0.0093 *n* = 135
Theta (–)	m = 0.0515 sd = 0.0121 *n* = 92	m = 0.0549 sd = 0.0117 *n* = 45	m = 0.0556 sd = 0.0121 *n* = 135

HR, heart rate; RMSSD, root mean square of successive differences between heart beats; BrR, breathing rate; SC, skin conductance; PD, pupil diameter; EBR, eye blink rate; EBD, eye blink duration; Alpha, relative EEG alpha power; Theta, relative EEG theta power; m, mean; sd, standard deviation; n, number of samples.

Q3) Do the Effects of Repetition Differ When the Participant Is Just Driving Compared to When Also Doing a Cognitive Task?

BrR and PD decreased significantly with increasing repetition in all task conditions, and the size of the decrease differed slightly between the task conditions for PD (demonstrated by the significant interaction effect between task and repetition; see **Table 3**). EBR, EBD, and Alpha showed an increasing trend with repetition in all task conditions, but only in EBD did these effects reach significance level in all task conditions. There were no significant interaction effects between task and repetition in EBR, EBD, or Alpha. For HR, the effect of repetition differed between task conditions. While HR decreased significantly with increasing repetition in the 1-back and 2-back tasks, there was no effect of repetition in the baseline condition.

Q4) How Do the Different Traffic Scenarios Affect Physiological Measures?

The mixed model ANOVAs revealed a significant effect of scenario on PD, EBR, and EBD (see **Table 3**). Their values for the Wind scenario consistently differed from those of the Intersection and Hidden Exit scenarios (except for EBR which did not differ significantly between Hidden Exit and Wind), while the latter two scenarios did not differ from each other.

Figure 2 shows the measures' continuous vectors for each scenario and task condition. For the Hidden Exit and Intersection scenarios, the measures are plotted in relation to the traffic environment: the x-axes represent distance driven. The plots are marked where the cognitive task begins and where the participants reach the hidden exit or intersection (depending on the scenario). There is no common position where the tasks end, since that depends on the vehicle speed. The average time between the task onset and the vehicle passing the hidden exit or intersection was ∼47 s. In the Wind scenario, the measures are plotted in relation to time, since the wind bursts were controlled by time, not position. Since the tasks' start and end depended on the vehicle speed (both the crosswind and the tasks began at a certain location in the simulated environment), there is neither a common task onset nor end in the plots. On average, the tasks began ∼10 s before the first large crosswind (first vertical line).

When the participants approached the hidden exit and intersection, EBR and EBD decreased consistently (except that the EBD decrease in the 2-back condition in the Intersection scenario did not reach significance), while SC and PD increased consistently. Some statistically significant results were found in HR, BrR, and Alpha, but they were inconsistent and small in relation to the signals' overall variability in the segments and are thus less likely to be actual and/or relevant effects.

Q5) Do the Effects of Traffic Scenario Differ When the Participant Is Just Driving Compared to When Also Doing a Cognitive Task?

The mixed model ANOVAs revealed no significant interaction effects between task and scenario in any measure (see **Table 3**). However, effect sizes appear to differ between task conditions when approaching and passing the hidden exit and intersection. The increase and decrease in the PD and EBD, respectively, were greater in the baseline and 1-back conditions compared to the 2-back condition. Note that to avoid excessive testing, no statistical testing has been done to compare these effect sizes.

DISCUSSION

The aim of this simulator study was to demonstrate and exemplify how the measurability of cognitive load can be improved by studying multiple mental responses, using multiple physiological measures and independent variables. We will refer to this as the *multidimensional approach* as it incorporates more than one mental response, measure, and independent variable. With this approach, the three aforementioned issues—(1) cognitive load consists of multiple mental responses, (2) cognitive load does not occur in isolation, and (3) physiological measures respond to multiple mental states—can be taken into

TABLE 3 | Results from Mixed Model ANOVAs of effects of task, repetition, and scenario for each measure.

	Main effect task	*Post-hoc* test: Bonferroni corrected *p*			Main effect repetition	Solution: estimate (se), *p*			Interaction effect repetition*task	Main effect scenario	*Post-hoc* test: Bonferroni corrected *p*			Interaction effect task*scenario
		Baseline vs. 1-back	**Baseline vs. 2-back**	**1-back vs. 2-back**		**Repetition* baseline**	**Repetition* 1-back**	**Repetition* 2-back**			**Hidden exit vs. Intersection**	**Intersection vs. Wind**	**Hidden exit vs. Wind**	
HR (beats/min)	$F_{(2,666)} = 51.48$, $p < 0.0001$	<0.0001	<0.0001	0.0002	$F_{(1,666)} = 36.84$, $p < 0.0001$	0.06 (0.15), $p = 0.71$	−0.97 (0.17), $p < 0.0001$	−0.92 (0.20), $p < 0.0001$	$F_{(2,666)} = 12.26$, $p < 0.0001$	$F_{(2,666)} = 1.57$, $p = 0.21$	0.34	1.0	0.41	$F_{(4,666)} = 1.69$, $p = 0.15$
RMSSD (–)	$F_{(2,666)} = 20.60$, $p < 0.0001$	<0.0001	<0.0001	<0.0001	$F_{(1,666)} = 4.63$, $p = 0.03$	0.0184 (0.0109), $p = 0.09$	0.0124 (0.0121), $p = 0.31$	0.0155 (0.0144), $p = 0.28$	$F_{(2,666)} = 0.07$, $p = 0.94$	$F_{(2,666)} = 0.08$, $p = 0.92$	1.0	1.0	1.0	$F_{(4,666)} = 1.33$, $p = 0.26$
BrR (breaths/min)	$F_{(2,454)} = 13.74$, $p < 0.0001$	<0.0001	<0.0001	0.0001	$F_{(1,454)} = 37.45$, $p < 0.0001$	−0.56 (0.11), $p < 0.0001$	−0.39 (0.12), $p = 0.002$	−0.42 (0.15), $p = 0.006$	$F_{(2,454)} = 0.56$, $p = 0.57$	$F_{(2,454)} = 2.56$, $p = 0.08$	0.78	0.81	0.07	$F_{(4,454)} = 0.57$, $p = 0.69$
SC (–)	$F_{(2,564)} = 13.46$, $p < 0.0001$	<0.0001	<0.0001	0.004	$F_{(1,564)} = 0.07$, $p = 0.79$	0.026 (0.013), $p = 0.05$	0.012 (0.015), $p = 0.40$	−0.031 (0.019), $p = 0.10$	$F_{(2,564)} = 3.07$, $p = 0.05$	$F_{(2,564)} = 4.07$, $p = 0.02$	1.0	0.10	0.02	$F_{(4,564)} = 2.57$, $p = 0.04$
PD (mm)	$F_{(2,398)} = 36.77$, $p < 0.0001$	<0.0001	<0.0001	<0.0001	$F_{(1,398)} = 124.78$, $p < 0.0001$	−0.081 (0.011), $p < 0.0001$	−0.112 (0.012), $p < 0.0001$	−0.046 (0.014), $p = 0.001$	$F_{(2,398)} = 6.29$, $p = 0.002$	$F_{(2,398)} = 238.19$, $p < 0.0001$	0.27	<0.0001	<0.0001	$F_{(4,398)} = 1.69$, $p = 0.15$
EBR (blinks/min)	$F_{(2,514)} = 11.61$, $p < 0.0001$	<0.0001	<0.0001	0.09	$F_{(1,514)} = 12.75$, $p = 0.0004$	1.32 (0.36), $p = 0.0003$	0.55 (0.40), $p = 0.16$	0.63 (0.47), $p = 0.18$	$F_{(2,514)} = 1.19$, $p = 0.30$	$F_{(2,514)} = 9.78$, $p < 0.0001$	0.13	<0.0001	0.04	$F_{(4,514)} = 0.94$, $p = 0.44$
EBD (ms)	$F_{(2,513)} = 1.42$, $p = 0.24$	0.12	0.10	0.002	$F_{(1,513)} = 118.13$, $p < 0.0001$	4.68 (0.84), $p < 0.0001$	7.68 (0.93), $p < 0.0001$	5.47 (1.09), $p < 0.0001$	$F_{(2,513)} = 2.89$, $p = 0.06$	$F_{(2,513)} = 13.45$, $p < 0.0001$	1.0	<0.0001	<0.0001	$F_{(4,513)} = 1.27$, $p = 0.28$
Alpha (–)	$F_{(2,239)} = 0.25$, $p = 0.78$		0.09		$F_{(1,239)} = 10.45$, $p = 0.001$	0.0008 (0.0004), $p = 0.02$	0.0005 (0.0005), $p = 0.35$	0.0008 (0.0003), $p = 0.004$	$F_{(2,239)} = 0.19$, $p = 0.82$	$F_{(2,239)} = 3.16$, $p = 0.04$	0.74			$F_{(3,239)} = 0.88$, $p = 0.45$
Theta (–)	$F_{(2,239)} = 0.62$, $p = 0.54$		<0.0001		$F_{(1,239)} = 6.14$, $p = 0.01$	−0.0011 (0.0006), $p = 0.04$	−0.0011 (0.0008), $p = 0.18$	−0.0004 (0.0005), $p = 0.40$	$F_{(2,239)} = 0.67$, $p = 0.51$	$F_{(2,239)} = 1.11$, $p = 0.33$	0.70			$F_{(3,239)} = 1.57$, $p = 0.20$

p-values are reported before Bonferroni correction for multiple tests. A significance level of 0.05 corresponds to 0.006 after Bonferroni correction. Green cells demark a $p < 0.006$. For consistency, results from post-hoc tests of differences between tasks and scenarios are included even where there is no significant main effect, although these effects are not color-coded. Since Alpha and Theta were only derived in Test Series 2, they lack the condition with the 1-back task in the Wind scenario. Thus, it was not possible to derive post-hoc results for all conditions for them. HR, heart rate; RMSSD, root mean square of successive differences between heart beats; BrR, breathing rate; SC, skin conductance; PD, pupil diameter; EBR, eye blink rate; EBD, eye blink duration; Alpha, relative EEG alpha power, and Theta, relative EEG theta power.

account. In this discussion section, the results from the five analysis questions will be interpreted using this multidimensional approach. Some alternative interpretations that overlook the issues will also be provided for the purpose of comparison. This alternative approach will be referred to as a *unidimensional approach*, since it views cognitive load as a unidimensional mental response.

Effects of Cognitive Tasks

The cognitive tasks had a significant effect on most physiological measures, namely HR, RMSSD, BrR, SC, PD, and EBR, in line with previous research (e.g., Mehler et al., 2009; Faure et al., 2016; Cegovnik et al., 2018). With a unidimensional approach that overlooks the three issues, one could stop the analysis here and conclude that these measures can therefore serve as indicators of cognitive load. We will, off course, not do that.

Note that contrary to what was expected, we saw no effect of the cognitive tasks on the EEG measures Theta and Alpha. One reason could be that we studied relative power instead of the more commonly used absolute power, which (as noted) sometimes show different effects (Wascher et al., 2018). The use of individually adapted frequency bands, instead of fixed bands as was used here, might also improve results (Klimesch, 1999). In addition, the equipment and methods used when deriving alpha and theta power differ a great deal between driving studies, making it difficult to compare results (Choi and Kim, 2018). Thus, the measures' limitations and possibilities in a driving context are still to be determined.

Including Effects of Repetition

When including the effects of repetition in the mental state assessment through a joint interpretation of HR, PD and RMSSD, the multidimensionality of cognitive load becomes evident. Recall that HR is a frequently used measure of cognitive load (Mehler et al., 2016; Hughes et al., 2019). In line with previous studies, increased cognitive task demand caused a stepwise increase in HR (e.g., Mehler et al., 2010). However, while the HR remained constant in the baseline condition, it decreased with repetition in both task conditions. In other words, the increase in HR caused by the cognitive tasks became smaller over time.

With a unidimensional approach in which cognitive load is viewed as a unidimensional construct whose level is reflected by HR, the decrease in HR would indicate that the level of cognitive load decreased over time. One could then assume, for example, that participants learned the tasks or gradually put less effort into doing them. However, the effects of repetition on PD and RMSSD speak against that interpretation. Increased cognitive task demand caused a stepwise increase in PD and a stepwise decrease in RMSSD, and, importantly, these effects were not attenuated with repetition. (To be precise, there was a significant interaction effect between task and repetition in PD, where the effect of 1-back attenuated slightly and the effect of 2-back increased slightly with increasing repetition. But for the sake of reasonable article length, we will not discuss this further.)

As explained, PD has a close neurological relation to cortical arousal and effort (van der Wel and van Steenbergen, 2018; Joshi and Gold, 2020), whereas the effect of cognitive demand on HR is more complex (Billman, 2013). In studies of mental workload, HR appears more driven by stress and negative emotion than cortical arousal, as the mentioned research on pilots have shown HR to be sensitive to workload alterations in real flying (Dussault et al., 2004) but not in simulated environments, where there is no physical risk involved (Dussault et al., 2005). Furthermore, it has been suggested that HRV has a closer relation to workload than HR (de Waard, 1996).

With the multidimensional approach that acknowledges that cognitive load has multiple components, the combined effects of the cognitive tasks and repetition suggest that different cognitive load components were differently affected by repetition. While the task-induced psychological stress decreased, the increase in cortical arousal and effort remained high throughout the experiment. This indicates that there was no learning or decrease in engagement effects after all.

Effects of repetition are not only seen as changes over time in the mental responses to the cognitive tasks, but also as changes in the participants' baseline state. As the baseline condition was repeated, EBR and EBD increased and BrR and PD decreased. With a unidimensional approach where only the level of cognitive load is of interest, effects of repetition (or time-on-task) are typically rendered insignificant as they are dealt with by employing a randomized or balanced test design. However, employing a multidimensional approach suggests incorporating these effects into the mental state interpretations rather than balancing them out.

The effects of repetition strongly suggest a decline in baseline level of arousal and attention. It seems that the drivers became less engaged with the driving task over time and became more fatigued. Also, HR and RMSSD in the baseline condition remained at relatively low and high levels, respectively, throughout the drive. It thus appears that the level of stress and effort related to the driving task was already relatively low at the first scenario repetition (recall that the participants had first practiced driving before the experimental session began), and that HR and RMSSD are not sensitive to further reductions in driving effort.

The participant's mental state can affect his/her physiological (Conway et al., 2013; Do et al., 2021), behavioral (Schoofs et al., 2008), and mental (Jimmieson et al., 2017; Hidalgo-Muñoz et al., 2018) responses to cognitive demand, so incorporating the baseline mental state when interpreting results improves the external validity and enables better comparisons between studies and environments.

Including Effects of Traffic Scenario

Here, effects on SC, PD, EBR, and EBD from the cognitive tasks and traffic scenarios are interpreted together. As the drivers approached the intersections and hidden exits, the SC and PD increased and the EBR and EBD decreased. With a unidimensional approach where only the level of cognitive load is assessed, these results appear conflicting. With a multidimensional approach that acknowledges multiple mental responses, these differences are instead informative. Remember that, unlike PD, which increases with increased attention regardless of attention modality, the eye blink measures decrease

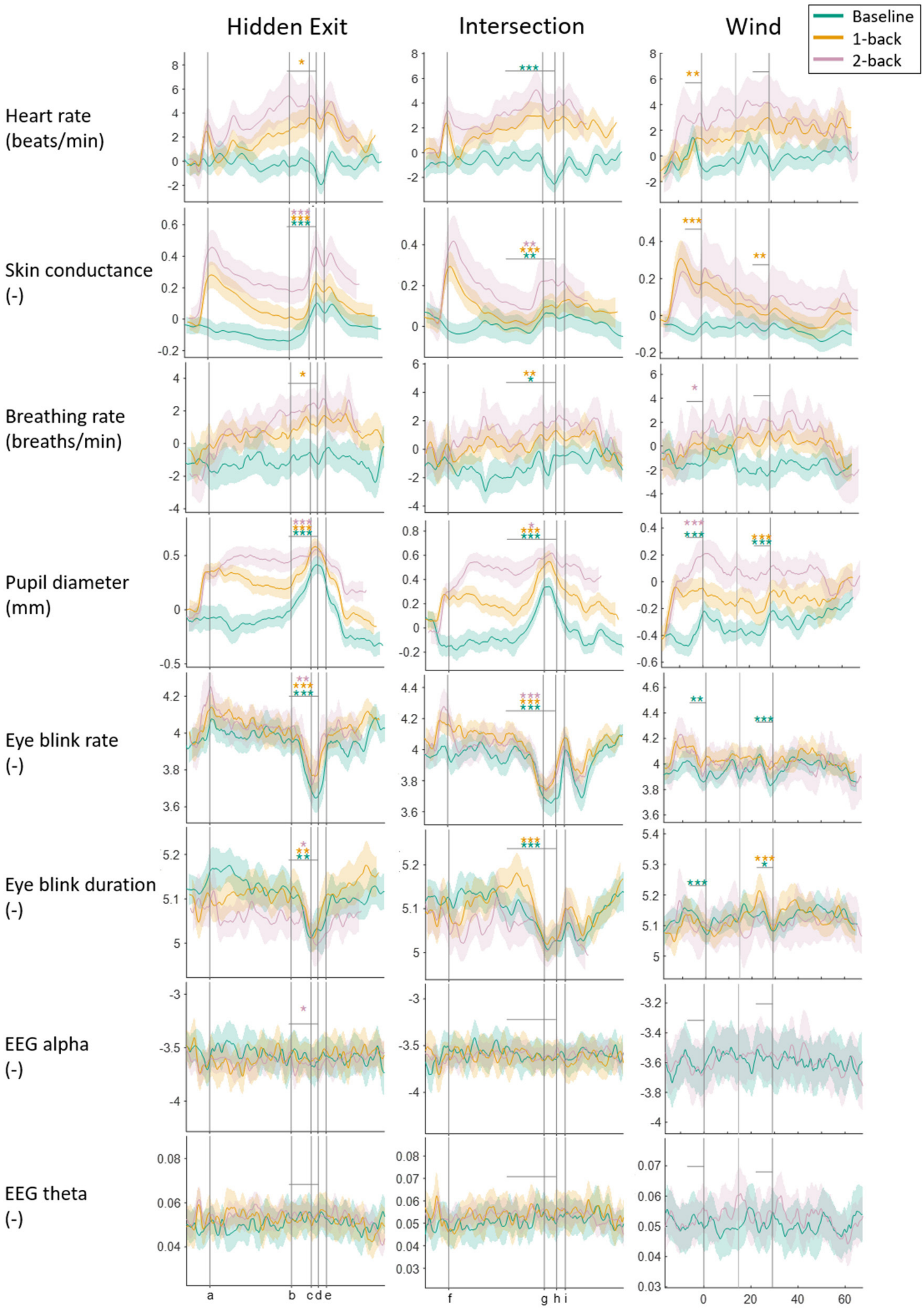


FIGURE 2 | Physiological measures plotted against the specific traffic events for each scenario and task condition. The thick colored lines are the means, and the shaded areas are the 95% confidence intervals. All measures, except Alpha and Theta, are normalized to reduce differences in absolute levels between participants.

(Continued)

FIGURE 2 | Curves are slightly smoothed to improve visibility. In the Hidden Exit scenario, the vertical lines show where the task begins (a); the hidden exit warning sign becomes visible (b); the warning sign is (c); the hidden exit becomes visible (d); and the exit is (e). In the Intersection scenario, vertical lines show where the task begins (f); the car approaching the intersection from the right becomes visible (g); the participant's car passes the oncoming bus (h); and the intersection is (i). In the Wind scenario, the three vertical lines mark the peaks of the three strongest crosswinds (the first and third are stronger than the second). Vertical gray lines show between which two points *t*-tests have been done, and the stars above the lines represent the results from the *t*-tests; $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

with increased visual attention, while non-visual attention (such as cognitive tasks) causes an increase in EBR (and sometimes also in EBD) (Recarte et al., 2008). It thus appears that when the participants approached the hidden exits and intersections, their cortical arousal increased due to increased visual attention, together with an increase in general arousal (as reflected in SC; Posada-Quintero and Chon, 2020).

In the case of the wind scenario, any effects of the crosswinds were less pronounced compared with the effects of the environmental demands in the other two scenarios. As noted, previous research employing crosswinds has suggested that the wind poses an additional cognitive demand (Medeiros-Ward et al., 2014), supported by physiological findings: a decrease in Alpha and an increase in Theta (Wascher et al., 2018). In contrast to Wascher et al.'s (2018) findings, there was no effect of crosswind on Alpha or Theta in our study. Recall, though, that there was no effect of the cognitive tasks (which we know cause an increase in cognitive load) on Alpha or Theta in our study, either. These EEG measures do thus not appear sensitive to cognitive load variations in this setting. However, the other physiological measures (which have proven sensitive to variations in several cognitive load components) improve our chances of registering a mental response, if there is one. The mixed model ANOVAs revealed no statistically significant effect of the crosswind on any of the measures, while the effects of individual wind bursts, visualized and statistically tested in **Figure 2**, showed mixed results. Since no correction for multiple testing has been done on these tests, they should be interpreted with extra consideration of response consistency to avoid type 1 errors. Only PD showed a fairly consistent effect of crosswinds with a significant effect in four out of six tests. It is thus plausible that the participants had brief increases in cortical arousal following the unpredictable crosswinds. But the combined results suggest that the crosswind posed only a very small cognitive load on the participants. Rather, the challenge of driving in a crosswind appears to have been dealt with quite automatically, without the driver having to assert much cognitive control (Schneider and Shiffrin, 1977). Although we employed similar crosswinds to those in Wascher et al.'s (2018) work, our study thus seems to have induced different mental responses. Although the reason is not known at this time, such differences in mental responses between studies could explain observed differences in behaviors between studies (see, e.g,. the different results in He et al., 2014, and Medeiros-Ward et al., 2014).

Implications of a Multidimensional Approach to Measuring Cognitive Load

The examples above demonstrate how the measurability of cognitive load can be improved by studying multiple mental responses using multiple physiological measures and independent variables. First, acknowledging that cognitive load is a multidimensional construct and measuring (some of) its components individually improves the construct validity of the study, compared to performing a unidimensional analysis (Strauss and Smith, 2009). It is clear from the examples above that several different mental responses occurred during the course of the experiment. For example, the psychological stress that the cognitive tasks gave rise to diminished over time, and visual attention increased with traffic complexity. Until we know how to weight different cognitive load components, it is thus not possible to assess the level of cognitive load on a unidimensional scale.

Having acknowledged that cognitive load is multidimensional and that its components need to be measured individually, the concurrent analysis of multiple physiological measures in relation to multiple independent variables improves the measures' diagnosticity. Making use of the different measures' similarities and differences makes it possible to look at multi-measure response patterns rather than single-measure responses. For example, changes in visual and non-visual attention could be distinguished from each other when PD and EBR or EBD were analyzed together.

At the same time, considering multi-measure response patterns instead of single-measure responses reduces the number of correlations to different mental states. The measurements' context dependence is thus reduced as fewer factors affect the same measurements. This means that the external validity is improved and the risk of making incorrect inferences from observed responses is reduced.

Most research seeking physiological indicators of cognitive load, especially if it employs machine learning, does indeed include multiple measures in the analyses (e.g., Putze et al., 2010; Murphey et al., 2019; Chihara et al., 2020). The use of multiple measures has also been encouraged for a long time (de Waard, 1996). Sometimes, the multiple measures are regarded as "backups" for each other (Tran et al., 2020) to mitigate issues with recording failures (Halverson et al., 2012) or individual response variability (Mehler et al., 2012), but often, multiple measure are indeed combined to improve classification accuracy (i.e., measurability) (e.g., Hogervorst et al., 2014; Prabhakar et al., 2020). However, if cognitive load is not acknowledged as a multidimensional construct, the issue of construct validity and the risk of making inaccurate inferences remain.

One could end up with measures that correlate only with certain cognitive load components that frequently occur in experiments (if that is where the training data are collected)—for example, measures reflecting psychological stress. There is a risk then that these cognitive load components do not occur as frequently in less controlled settings, such as self-initiated cognitive tasks in real-life driving (de Waard, 1996). Consequently, such a measure might fail to detect cognitive load under less stressful circumstances, even if the loading on other cognitive load components is significant.

By measuring and studying cognitive load components separately, researchers can assess the components' individual and combined effects. They can, for example, explore the effects of cognitive effort and psychological stress, separately and together, on driver behavior and traffic safety. Car manufacturers can then use the information gained to prioritize those mental states which are most relevant to detect in Driver Monitoring Systems (DMS), for example. However, there are several challenges when going from group-level studies to continuous monitoring of drivers' mental states.

One great challenge for DMS systems is that between-subject variance in physiological responses to cognitive load (and other mental states) is large (Mehler et al., 2012). Individualized algorithms have therefore been suggested for accurate tracking (Noh et al., 2021). One advantage of tracking multiple mental responses is that the between-subject differences in the physiological responses to changes in individual cognitive load components should be smaller than the differences in physiological responses to cognitive load as a whole (i.e., when it is studied as a unidimensional construct). This is due to the fact that not all drivers have the same mental responses, such as increased psychological stress, during increased cognitive demands (Szalma, 2008). DMS development might thus be somewhat less complicated if cognitive load assessment is made multidimensional.

Still, variability will remain an issue since not all drivers have the same physiological responses to the same mental state changes (e.g., not all individuals display frontal-midline theta activity; Mitchell et al., 2008). While some of this variability could possibly be reduced by breaking down mental responses further, that may render the analysis overly complex. Also, not all cognitive functions can be continuously measured in car drivers. In the end, the appropriate level of detail is one that enables researchers and car manufacturers to understand and, when needed, mitigate any negative effects of cognitive tasks on traffic safety, without making the mental state assessment overly complicated.

It should also be noted that effects seen on a group-level are not necessarily detectable at an individual level because of the multiple factors concurrently influencing the physiological measures. This is especially true where effect sizes are small. For example, the size of mentally driven changes in the PD are typically below 0.5 mm (Beatty, 1982), while alterations in lightening can change the PD several millimeters (Winn et al., 1994).

Study Limitations

This experiment was designed for many purposes (Nilsson et al., 2018; see also Nilsson et al., 2020), which limited the design possibilities somewhat. Priority was given to achieving a realistic driving task with a low level of interference, which prevented the use of subjective estimates while driving.

The aim of this study was to use multiple physiological measures and independent variables to assess multiple mental responses and, by that, improve the cognitive load measurability. However, only a limited set of physiological measures was included. The measurability can likely be improved using more, and/or other, physiological signals and measures. It may also be that some of the measures are less sensitive in other environments, such as real driving.

The physiological data came from a fairly homogenous group of participants. The variability in responses to the experimental manipulations may therefore be smaller than would have been the case in a more heterogenous group.

Finally, multiple statistical tests were conducted (which is hard to avoid when interpreting multiple measures and independent variables). Bonferroni corrections were made on the ANOVA results to decrease the risk of type 1 errors, while no correction for multiple tests were made for the *t*-test results to avoid inflating the risk of type 2 errors and disregarding actual effects (Forstmeier et al., 2017). To deal with the increased risk of type 1 errors, consistency in results and effect sizes were considered in the interpretations. Still, effects seen in the continuous plots and *t*-tests should be considered exploratory and in need of verification in future studies.

Overall, as the complex relationships between coexisting mental states and physiological responses are still largely unknown, the inferences we made from the physiological measures are, to some extent, speculative. There are also no established "ground truth" measures of mental states to validate our interpretations. Using non-physiological measures, such as questionnaires and performance metrics, could improve the validity of the interpretations (Hancock and Matthews, 2019), although all measures have their own limitations. For example, questionnaires can interfere with the driving task and make it less realistic (O'Donnell and Eggemeier, 1986); people are sometimes not very good at self-assessing their mental state (Schmidt et al., 2009); and performance measures typically have a limited range of sensitivity, since performance can be modulated with effort (Reimer et al., 2012).

CONCLUSIONS

In conclusion, when cognitive load is understood as a multidimensional construct, and (some of) its components are assessed separately using multiple physiological measures studied in relation to multiple independent variables, its measurability can be improved in several ways. For one, the construct validity of cognitive load is improved, which facilitates more accurate and useful result interpretations. Also, studied together and related to multiple mental states, the measures are more diagnostic, in that they are better able to distinguish between changes in different cognitive load components. With multiple measures, multi-measure response patterns can be analyzed instead of single-measure responses. Since the patterns correlate with fewer mental responses, the measurements' external validity is also improved, and the risk of making incorrect inferences from observed responses is reduced.

Improved measurability of cognitive load has the potential to enable more detailed and accurate inferences regarding the effects of cognitive task execution in less controlled settings. As a result, the effects of cognitive load on traffic safety can be better understood and more effectively mitigated.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because they are proprietary. Further inquiries about the dataset can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Regionala etikprövningsnämnden, Linköping. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

EN and BS designed the study and contributed to the data acquisition and data curation. EN derived the physiological measures, performed the analyses and visualizations, and wrote the original draft manuscript. EN, JB, ML, and GM performed iterative reviews and edits. JB and ML also supervised EN. All authors contributed to the article and approved the submitted version.

FUNDING

The experiments presented herein were part of the Vehicle Driver Monitoring (VDM) project, sponsored by Vinnova, the Swedish governmental agency for innovation systems. The funding agency was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication. The final analyses and writing of this article were financed by Volvo Car Corporation. Only the authors analyzed, interpreted the data, and wrote the paper.

ACKNOWLEDGMENTS

We would like to thank Per Lindén, Pär Gustavsson, and Louise Walletun for their contribution to data collection and data preparation and Kristina Mayberry for her language review.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnrgo.2022.787295/full#supplementary-material

REFERENCES

Abd Rahman, N. I., Md Dawal, S. Z., and Yusoff, N. (2020). Driving mental workload and performance of ageing drivers. *Transp. Res. F Traffic Psychol. Behav.* 69, 265–285. doi: 10.1016/j.trf.2020.01.019

Ahlström, C., Anund, A., Fors, C., and Åkerstedt, T. (2018). Effects of the road environment on the development of driver sleepiness in young male drivers. *Accid. Anal. Prevent.* 112, 127–134. doi: 10.1016/j.aap.2018.01.012

Ahlström, C., Zemblys, R., Jansson, H., Forsberg, C., Karlsson, J., and Anund, A. (2021). Effects of partially automated driving on the development of driver sleepiness. *Accid. Anal. Prevent.* 153:106058. doi: 10.1016/j.aap.2021.106058

Almahasneh, H., Chooi, W. T., Kamel, N., and Malik, A. S. (2014). Deep in thought while driving: an EEG study on drivers' cognitive distraction. *Transp. Res. F Traffic Psychol. Behav.* 26, 218–226. doi: 10.1016/j.trf.2014.08.001

Aston-Jones, G., and Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annu. Rev. Neurosci.* 28, 403–450. doi: 10.1146/annurev.neuro.28.061604.135709

Backs, R. W., and Seljos, K. A. (1994). Metabolic and cardiorespiratory measures of mental effort: the effects of level of difficulty in a working memory task. *Int. J. Psychophysiol.* 16, 57–68. doi: 10.1016/0167-8760(94)90042-6

Bafna, T., and Hansen, J. P. (2021). Mental fatigue measurement using eye metrics: a systematic literature review. *Psychophysiology* 6:e13828. doi: 10.1111/psyp.13828

Baker, L. B. (2019). Physiology of sweat gland function: the roles of sweating and sweat composition in human health. *Temperature* 6, 211–259. doi: 10.1080/23328940.2019.1632145

Balters, S., and Steinert, M. (2015). Capturing emotion reactivity through physiology measurement as a foundation for affective engineering in engineering design science and engineering practices. *J. Intelligent Manufact.* 28:1585–607. doi: 10.1007/s10845-015-1145-2

Barzegaran, E., Vildavski, V. Y., and Knyazeva, M. G. (2017). Fine structure of posterior alpha rhythm in human EEG: frequency components, their cortical sources, and temporal behavior. *Sci. Rep.* 7:8249. doi: 10.1038/s41598-017-08421-z

Baumann, M., Petzoldt, T., Groenewoud, C., Hogema, J., and Krems, J. F. (2008). "The effect of cognitive tasks on predicting events in traffic," in *Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems* (Lyon), 3–12.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull.* 91, 276–292. doi: 10.1037/0033-2909.91.2.276

Beggiato, M., Hartwich, F., and Krems, J. (2019). Physiological correlates of discomfort in automated driving. *Transp. Res. F Traffic Psychol. Behav.* 66, 445–458. doi: 10.1016/j.trf.2019.09.018

Belyusar, D., Mehler, B., Solovey, E., and Reimer, B. (2015). Impact of repeated exposure to a multilevel working memory task on physiological arousal and driving performance. *Transp. Res. Rec.* 2518, 46–53. doi: 10.3141/2518-06

Benarroch, E. E. (2011). "Central autonomic control," in *Primer on the Autonomic Nervous System* (3rd ed.), eds P. A. Low, R. J. Polinsky, D. Robertson, and J. F. R. Paton (London; Waltham, MA; San Diego, CA: Elsevier Science & Technology), 9–12. doi: 10.1016/B978-0-12-386525-0.00002-0

Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., and Montanari, R. (2011). Driver workload and eye blink duration. *Transp. Res. F Traffic Psychol. Behav.* 14, 199–208. doi: 10.1016/j.trf.2010.12.001

Benwell, C. S. Y., London, R. E., Tagliabue, C. F., Veniero, D., Gross, J., Keitel, C., et al. (2019). Frequency and power of human alpha oscillations drift systematically with time-on-task. *Neuroimage* 192, 101–114. doi: 10.1016/j.neuroimage.2019.02.067

Berberian, B., Somon, B., Saha,ï, A., and Gouraud, J. (2017). The out-of-the-loop Brain: a neuroergonomic approach of the human automation interaction. *Annu. Rev. Control* 44, 303–315. doi: 10.1016/j.arcontrol.2017.09.010

Billman, G. E. (2013). The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. *Front. Physiol.* 4:26. doi: 10.3389/fphys.2013.00026

Borghini, G., Aricò, P., Graziani, I., Salinari, S., Sun, Y., Taya, F., et al. (2016). Quantitative assessment of the training improvement in a motor-cognitive task by using EEG, ECG and EOG signals. *Brain Topogr.* 29, 149–161. doi: 10.1007/s10548-015-0425-7

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., and Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neurosci. Biobehav. Rev.* 44, 58–75. doi: 10.1016/j.neubiorev.2012.10.003

Bradley, M. M. (2009). Natural selective attention: orienting and emotion. *Psychophysiology* 46, 1–11. doi: 10.1111/j.1469-8986.2008.00702.x

Braithwaite, J. J., Watson, D. G., Jones, R., and Rowe, M. (2015). *A Guide for Analysing Electrodermal Activity (EDA) and Skin Conductance Responses (SCRs) for Psychological Experiments (Revised version: 2.0)*. Birmingham: Selective Attention & Awareness Laboratory (SAAL) Behavioural Brain Sciences Centre, University of Birmingham, UK.

Brouwer, A. M., Hogervorst, M. A., Holewijn, M., and van Erp, J. B. F. (2014). Evidence for effects of task difficulty but not learning on neurophysiological variables associated with effort. *Int. J. Psychophysiol.* 93, 242–252. doi: 10.1016/j.ijpsycho.2014.05.004

Buendia, R., Forcolin, F., Karlsson, J., Arne Sjöqvist, B., Anund, A., and Candefjord, S. (2019). Deriving heart rate variability indices from cardiac monitoring—an indicator of driver sleepiness. *Traffic Inj. Prev.* 20, 249–254. doi: 10.1080/15389588.2018.1548766

Caird, J. K., Johnston, K. A., Willness, C. R., Asbridge, M., and Steel, P. (2014). A meta-analysis of the effects of texting on driving. *Accid. Anal. Prevent.* 71, 311–318. doi: 10.1016/j.aap.2014.06.005

Campbell, D. T., and Stanley, J. C. (1963). "Experimental and quasi-experimental designs for research," in *Handbook of Research on Teaching*, ed N. L. Gage (Chicago, IL: Rand McNally & Company). Available online at: https://www.jameslindlibrary.org/campbell-dt-stanley-jc-1963/

Cardona, G., García, C., Serés, C., Vilaseca, M., and Gispets, J. (2011). Blink rate, blink amplitude, and tear film integrity during dynamic visual display terminal tasks. *Curr. Eye Res.* 36, 190–197. doi: 10.3109/02713683.2010.544442

Cardona, G., and Quevedo, N. (2014). Blinking and driving: the influence of saccades and cognitive workload. *Curr. Eye Res.* 39, 239–244. doi: 10.3109/02713683.2013.841256

Cavanagh, J. F., and Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* 18, 414–421. doi: 10.1016/j.tics.2014.04.012

Cavanagh, J. F., and Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. *J. Physiol.* 109, 3–15. doi: 10.1016/j.jphysparis.2014.04.003

Cegovnik, T., Stojmenova, K., Jakus, G., and Sodnik, J. (2018). An analysis of the suitability of a low-cost eye tracker for assessing the cognitive load of drivers. *Appl. Ergon.* 68, 1–11. doi: 10.1016/j.apergo.2017.10.011

Chihara, T., Kobayashi, F., and Sakamoto, J. (2020). Evaluation of mental workload during automobile driving using one-class support vector machine with eye movement data. *Appl. Ergon.* 89:103201. doi: 10.1016/j.apergo.2020.103201

Choi, J. W., and Kim, K. H. (2018). *Computational EEG Analysis*, ed C.-H. Im (Singapore: Springer).

Clayton, M. S., Yeung, N., and Cohen Kadosh, R. (2015). The roles of cortical oscillations in sustained attention. *Trends Cogn. Sci.* 19, 188–195. doi: 10.1016/j.tics.2015.02.004

Compton, R. J., Gearinger, D., and Wild, H. (2019). The wandering mind oscillates: EEG alpha power is enhanced during moments of mind-wandering. *Cogn. Affect. Behav. Neurosci.* 19, 1184–1191. doi: 10.3758/s13415-019-00745-9

Conway, D., Dick, I., Li, Z., Wang, Y., and Chen, F. (2013). "The effect of stress on cognitive load measurement," in *Human-Computer Interaction - INTERACT 2013. Lecture Notes in Computer Science*, Vol. 8120, eds P. Kotzé, G. Marsden, G. Lindgaard, J. Wesson, and M. Winckler (Berlin, Heidelberg: Springer), 659–666. doi: 10.1007/978-3-642-40498-6_58

Cooke, A. (2013). Readying the head and steadying the heart: a review of cortical and cardiac studies of preparation for action in sport. *Int. Rev. Sport Exerc. Psychol.* 6, 122–138. doi: 10.1080/1750984X.2012.724438

Cori, J. M., Anderson, C., Shekari Soleimanloo, S., Jackson, M. L., and Howard, M. E. (2019). Narrative review: do spontaneous eye blink parameters provide a useful assessment of state drowsiness? *Sleep Med. Rev.* 45, 95–104. doi: 10.1016/j.smrv.2019.03.004

Critchley, H. D., Eccles, J., and Garfinkel, S. N. (2013). "Interaction between cognition, emotion, and the autonomic nervous system," in *Handbook of Clinical Neurology*, Vol. 117, eds R. M. Buijs and D. F. Swaab (Amsterdam: Elsevier), 59–77. doi: 10.1016/B978-0-444-53491-0.00006-7

Cruz, A. A. V., Garcia, D. M., Pinto, C. T., and Cechetti, S. P. (2011). Spontaneous eyeblink activity. *Ocular Surface* 9, 29–41. doi: 10.1016/S1542-0124(11)70007-6

Cumming, G., and Finch, S. (2005). Inference by eye confidence intervals and how to read pictures of data. *Am. Psychol.* 60, 170–180. doi: 10.1037/0003-066X.60.2.170

Dawson, M. E., Schell, A. M., and Filion, D. L. (2016). "The electrodermal system," in *Handbook of Psychophysiology*, 4th ed, eds J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson (Cambridge: Cambridge University Press), 217–243. doi: 10.1017/9781107415782.010

de Waard, D. (1996). *The Measurement of Drivers' Mental Workload* (Doctoral dissertation, University of Groningen). Retrieved from: http://dissertations.ub.rug.nl/FILES/faculties/ppsw/1996/d.de.waard/09_thesis.pdf (accessed May 19, 2021).

Dehais, F., Lafont, A., Roy, R., and Fairclough, S. (2020). A neuroergonomics approach to mental workload, engagement and human performance. *Front. Neurosci.* 14:268. doi: 10.3389/fnins.2020.00268

Del Negro, C. A., Funk, G. D., and Feldman, J. L. (2018). Breathing matters. *Nat. Rev. Neurosci.* 19, 351–367. doi: 10.1038/s41583-018-0003-6

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Desmond, P. A., and Hancock, P. A. (2001). "Active and passive fatigue states," in *Stress, Workload, and Fatigue*, eds P. A. Hancock, and P. A. Desmond (Mahwah, NJ: Lawrence Erlbaum Associates), 455–465. doi: 10.1201/b12791-3.1

Di Flumeri, G., Borghini, G., Aricò, P., Sciaraffa, N., Lanzi, P., Pozzi, S., et al. (2018). EEG-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings. *Front. Hum. Neurosci.* 12:509. doi: 10.3389/fnhum.2018.00509

Diaz-Piedra, C., Sebastián, M. V., and Di Stasi, L. L. (2020). EEG theta power activity reflects workload among army combat drivers: an experimental study. *Brain Sci.* 10:199. doi: 10.3390/brainsci10040199

Do, T. T. N., Wang, Y. K., and Lin, C. T. (2021). Increase in brain effective connectivity in multitasking but not in a high-fatigue state. *IEEE Trans. Cogn. Dev. Syst.* 13, 566–574. doi: 10.1109/TCDS.2020.2990898

Dussault, C., Jouanin, J. C., and Guezennec, C. Y. (2004). EEG and ECG changes during selected flight sequences. *Aviat. Space Environ. Med.* 75, 889–897.

Dussault, C., Jouanin, J. C., Philippe, M., and Guezennec, C. Y. (2005). EEG and ECG changes during simulator operation reflect mental workload and vigilance. *Aviat. Space Environ. Med.* 76, 344–351.

Eckstein, M. K., Guerra-Carrillo, B., Miller Singley, A. T., and Bunge, S. A. (2017). Beyond eye gaze: what else can eyetracking reveal about cognition and cognitive development? *Dev. Cogn. Neurosci.* 25, 69–91. doi: 10.1016/j.dcn.2016.11.001

Eilers, P. H. C. (2003). A perfect smoother. *Anal. Chem.* 75, 3631–3636. doi: 10.1021/ac034173t

Engström, J., Aust, M. L., and Viström, M. (2010). Effects of working memory load and repeated scenario exposure on emergency braking performance. *Hum. Factors* 52, 551–559. doi: 10.1177/0018720810381072

Engström, J., Markkula, G., Victor, T., and Merat, N. (2017). Effects of cognitive load on driving performance: the cognitive control hypothesis. *Hum. Factors* 59, 734–764. doi: 10.1177/0018720817690639

Engström, J., Monk, C., Hanowski, R., Horrey, W., Lee, J., McGehee, D., et al. (2013). *A Conceptual Framework and Taxonomy for Understanding and Categorizing Driver Inattention*. (Brussels: European Commission, Directorate General for Communications Networks, Content & Technology).

Faure, V., Lobjois, R., and Benguigui, N. (2016). The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior. *Transp. Res. F Traffic Psychol. Behav.* 40, 78–90. doi: 10.1016/j.trf.2016.04.007

Fitch, G. M., Soccolich, S. A, Guo, F., McClafferty, J., Fang, Y., Olson, R. L., et al. (2013). *The Impact of Hand-Held And Hands-Free Cell Phone Use on*

Driving Performance and Safety-Critical Event Risk, Report No: DOT HS 811 757. (Washington, DC: National Highway Traffic Safety Administration).

Fogarty, C., and Stern, J. A. (1989). Eye movements and blinks: their relationship to higher cognitive processes. *Int. J. Psychophysiol.* 8, 35–42. doi: 10.1016/0167-8760(89)90017-2

Forstmeier, W., Wagenmakers, E. J., and Parker, T. H. (2017). Detecting and avoiding likely false-positive findings – a practical guide. *Biol. Rev.* 92, 1941–1968. doi: 10.1111/brv.12315

Foy, H. J., and Chapman, P. (2018). Mental workload is reflected in driver behaviour, physiology, eye movements and prefrontal cortex activation. *Appl. Ergon.* 73, 90–99. doi: 10.1016/j.apergo.2018.06.006

Geršak, G. (2020). Electrodermal activity - a beginner's guide. *Electrotechn. Rev.* 87, 175–182.

Gomez, P., von Gunten, A., and Danuser, B. (2016). Autonomic nervous system reactivity within the valence–arousal affective space: modulation by sex and age. *Int. J. Psychophysiol.* 109, 51–62. doi: 10.1016/j.ijpsycho.2016.10.002

Grassmann, M., Vlemincx, E., von Leupoldt, A., Mittelstädt, J. M., and Van Den Bergh, O. (2016). Respiratory changes in response to cognitive load: a systematic review. *Neural Plast.* 2016:8146809. doi: 10.1155/2016/8146809

Grassmann, M., Vlemincx, E., von Leupoldt, A., and Van den Bergh, O. (2017). Individual differences in cardiorespiratory measures of mental workload: an investigation of negative affectivity and cognitive avoidant coping in pilot candidates. *Appl. Ergon.* 59, 274–282. doi: 10.1016/j.apergo.2016.09.006

Grueschow, M., Kleim, B., and Ruff, C. C. (2020). Role of the locus coeruleus arousal system in cognitive control. *J. Neuroendocrinol.* 32, 1–11. doi: 10.1111/jne.12890

Halgren, M., Ulbert, I., Bastuji, H., Fab,ó, D., Eross, L., Rey, M., et al. (2019). The generation and propagation of the human alpha rhythm. *Proc. Natl. Acad. Sci. U.S.A.* 116, 23772–23782. doi: 10.1073/pnas.1913092116

Halverson, T., Estepp, J., Christensen, J., and Monnin, J. (2012). "Classifying workload with eye movements in a complex task," in *Proceedings of the Human Factors and Ergonomics Society* (Boston, MA), 168–172. doi: 10.1177/1071181312561012

Hammel, K. R., Fisher, D. L., and Pradhan, A. K. (2002). "Verbal and spatial loading effects on eye movements in driving simulators: a comparison to real world driving," *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (Baltimore, MD), 2174–2178. doi: 10.1177/154193120204602615

Hancock, P. A., and Matthews, G. (2019). Workload and performance: associations, insensitivities, and dissociations. *Hum. Factors* 61, 374–392. doi: 10.1177/0018720818809590

He, D., Donmez, B., Liu, C. C., and Plataniotis, K. N. (2019). High cognitive load assessment in drivers through wireless electroencephalography and the validation of a modified N -back task. *IEEE Trans. Human Machine Syst.* 49, 362–371. doi: 10.1109/THMS.2019.2917194

He, J., McCarley, J. S., and Kramer, A. F. (2014). Lane keeping under cognitive load: performance changes and mechanisms. *Hum. Factors* 56, 414–426. doi: 10.1177/0018720813485978

Heine, T., Lenis, G., Reichensperger, P., Beran, T., Doessel, O., and Deml, B. (2017). Electrocardiographic features for the measurement of drivers' mental workload. *Appl. Ergon.* 61, 31–43. doi: 10.1016/j.apergo.2016.12.015

Helfrich, R. F., and Knight, R. T. (2016). Oscillatory dynamics of prefrontal cognitive control. *Trends Cogn. Sci.* 20, 916–930. doi: 10.1016/j.tics.2016.09.007

Hidalgo-Muñoz, A. R., Béquet, A. J., Astier-Juvenon, M., Pépin, G., Fort, A., Jallais, C., et al. (2019). Respiration and heart rate modulation due to competing cognitive tasks while driving. *Front. Hum. Neurosci.* 12:525. doi: 10.3389/fnhum.2018.00525

Hidalgo-Muñoz, A. R., Mouratille, D., Matton, N., Causse, M., Rouillard, Y., and El-Yagoubi, R. (2018). Cardiovascular correlates of emotional state, cognitive workload and time-on-task effect during a realistic flight simulation. *Int. J. Psychophysiol.* 128, 62–69. doi: 10.1016/j.ijpsycho.2018.04.002

Hogervorst, M. A., Brouwer, A. M., and van Erp, J. B. F. (2014). Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload. *Front. Neurosci.* 8:322. doi: 10.3389/fnins.2014.00322

Homma, I., and Masaoka, Y. (2008). Breathing rhythms and emotions. *Exp. Physiol.* 93, 1011–1021. doi: 10.1113/expphysiol.2008.042424

Horne, J., and Reyner, L. (1999). Vehicle accidents related to sleep: a review. *Occup. Environ. Med.* 56, 289–294. doi: 10.1136/oem.56.5.289

Hu, X., and Lodewijks, G. (2020). Detecting fatigue in car drivers and aircraft pilots by using non-invasive measures: the value of differentiation of sleepiness and mental fatigue. *J. Safety Res.* 72, 173–187. doi: 10.1016/j.jsr.2019.12.015

Hughes, A. M., Hancock, G. M., Marlow, S. L., Stowers, K., and Salas, E. (2019). Cardiac measures of cognitive workload: a meta-analysis. *Hum. Factors* 61, 393–414. doi: 10.1177/0018720819830553

International Organization of Standardization (2014). *Road Vehicles - Measurement of Driver Visual Behaviour With Respect to Transport Information and Control Systems Part 1 : Definitions and Parameters*.

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., and Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory* 18, 394–412. doi: 10.1080/09658211003702171

Jammes, B., Sharabty, H., and Esteve, D. (2008). Automatic EOG analysis: a first step toward automatic drowsiness scoring during wake-sleep transitions. *Somnologie* 12, 227–232. doi: 10.1007/s11818-008-0351-y

Jiménez-Buedo, M., and Russo, F. (2021). Experimental practices and objectivity in the social sciences: re-embedding construct validity in the internal–external validity distinction. *Synthese* 199, 9549–9579. doi: 10.1007/s11229-021-03215-3

Jimmieson, N. L., Tucker, M. K., and Walsh, A. J. (2017). Interaction effects among multiple job demands: an examination of healthcare workers across different contexts. *Anxiety Stress Coping* 30, 317–332. doi: 10.1080/10615806.2016.1229471

Jones, S. R. (2016). When brain rhythms aren't 'rhythmic': implication for their mechanisms and meaning. *Curr. Opin. Neurobiol.* 40, 72–80. doi: 10.1016/j.conb.2016.06.010

Jongkees, B. J., and Colzato, L. S. (2016). Spontaneous eye blink rate as predictor of dopamine-related cognitive function—a review. *Neurosci. Biobehav. Rev.* 71, 58–82. doi: 10.1016/j.neubiorev.2016.08.020

Joshi, S., and Gold, J. I. (2020). Pupil size as a window on neural substrates of cognition. *Trends Cogn. Sci.* 24, 466–480. doi: 10.1016/j.tics.2020.03.005

Karlsson, M., Hörnsten, R., Rydberg, A., and Wiklund, U. (2012). Automatic filtering of outliers in RR intervals before analysis of heart rate variability in Holter recordings: a comparison with carefully edited data. *Biomed. Eng. Online* 11:21. doi: 10.1186/1475-925X-11-2

Kelsey, R. M., Soderlund, K., and Arthur, C. M. (2004). Cardiovascular reactivity and adaptation to recurrent psychological stress : replication and extension. *Psychophysiology* 41, 924–934. doi: 10.1111/j.1469-8986.2004.00245.x

Kiashari, S. E. H., Nahvi, A., Bakhoda, H., Homayounfard, A., and Tashakori, M. (2020). Evaluation of driver drowsiness using respiration analysis by thermal imaging on a driving simulator. *Multimed. Tools Appl.* 79, 17793–17815. doi: 10.1007/s11042-020-08696-x

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* 29, 169–195. doi: 10.1016/S0165-0173(98)00056-3

Klingner, J. (2010). *Measuring Cognitive Load During Visual Tasks by Combining Pupillometry and Eye Tracking*. Stanford, CA: Stanford University.

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biol. Psychol.* 84, 394–421. doi: 10.1016/j.biopsycho.2010.03.010

Laborde, S., Mosley, E., and Thayer, J. F. (2017). Heart rate variability and cardiac vagal tone in psychophysiological research - recommendations for experiment planning, data analysis, and data reporting. *Front. Psychol.* 8:213. doi: 10.3389/fpsyg.2017.00213

Larsen, R. S., and Waters, J. (2018). Neuromodulatory correlates of pupil dilation. *Front. Neural Circuits* 12:21. doi: 10.3389/fncir.2018.00021

Li, Z., Jiao, K., Chen, M., Yang, Y., Wang, C., and Qi, S. (2002). Spectral analysis of heart rate variability as a quantitative indicator of driver mental fatigue. *SAE Technical Paper Series*. 2002-01-0090, 1–5. doi: 10.4271/2002-01-0090

Liang, Y., and Lee, J. D. (2010). Combining cognitive and visual distraction: less than the sum of its parts. *Accid. Anal. Prevent.* 42, 881–890. doi: 10.1016/j.aap.2009.05.001

Lobjois, R., Faure, V., Désir,é, L., and Benguigui, N. (2021). Behavioral and workload measures in real and simulated driving: do they tell us the same thing about the validity of driving simulation? *Saf. Sci.* 134, 1–10. doi: 10.1016/j.ssci.2020.105046

Loeches De La Fuente, H., Berthelon, C., Fort, A., Etienne, V., De Weser, M., Ambeck, J., et al. (2019). Electrophysiological and performance variations following driving events involving an increase in mental workload. *Europ. Transp. Res. Rev.* 11:42. doi: 10.1186/s12544-019-0379-z

Lohani, M., Payne, B. R., and Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Front. Hum. Neurosci.* 13:57. doi: 10.3389/fnhum.2019.00057

Magliacano, A., Fiorenza, S., Estraneo, A., and Trojano, L. (2020). Eye blink rate increases as a function of cognitive load during an auditory oddball paradigm. *Neurosci. Lett.* 736:135293. doi: 10.1016/j.neulet.2020.135293

Makeig, S., Bell, A. J., Jung, T.-P., and Sejnowski, T. J. (1996). "Independent component analysis of electroencephalographic data," in *Advances in Neural Information Processing Systems*, eds D. Touretzky, M. Mozer, and M. Hasselmo (Cambridge, MA: MIT Press), 145–151.

Marzano, C., Fratello, F., Moroni, F., Pellicciari, M. C., Curcio, G., Ferrara, M., et al. (2007). Slow eye movements and subjective estimates of sleepiness predict EEG power changes during sleep deprivation. *Sleep* 30, 610–616. doi: 10.1093/sleep/30.5.610

Matthews, G., Reinerman-Jones, L. E., Barber, D. J., and Abich, J. (2015). The psychometrics of mental workload: multiple measures are sensitive but divergent. *Hum. Factors* 57, 125–143. doi: 10.1177/0018720814539505

Matthews, G., Wohleber, R., Lin, J., Funke, G., and Neubauer, C. (2019). "Monitoring task fatigue in contemporary and future vehicles: A review," in: *Advances in Intelligent Systems and Computing*, Vol. 780, ed D. Cassenti (Cham: Springer), 101–112. doi: 10.1007/978-3-319-94223-0_10

Matuz, A., van der Linden, D., Kisander, Z., Hernádi, I., Kázmér, K., and Csathó, Á. (2021). Enhanced cardiac vagal tone in mental fatigue: analysis of heart rate variability in Time-on-Task, recovery, and reactivity. *PLoS ONE* 16:e0238670. doi: 10.1371/journal.pone.0238670

Medeiros-Ward, N., Cooper, J. M., and Strayer, D. L. (2014). Hierarchical control and driving. *J. Exp. Psychol.* 143, 953–958. doi: 10.1037/a0035097

Mehler, B., Kidd, D., Reimer, B., Reagan, I., Dobres, J., and McCartt, A. (2016). Multi-modal assessment of on-road demand of voice and manual phone calling and voice navigation entry across two embedded vehicle systems. *Ergonomics* 59, 344–367. doi: 10.1080/00140139.2015.1081412

Mehler, B., Reimer, B., and Coughlin, J. F. (2010). Physiological reactivity to graded levels of cognitive workload across three age groups: an on-road evaluation. *Proc. Human Factors Ergonom. Soc. Annual Meet.* 54, 2062–2066. doi: 10.1177/154193121005402409

Mehler, B., Reimer, B., and Coughlin, J. F. (2012). Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: an on-road study across three age groups. *Hum. Factors* 54, 396–412. doi: 10.1177/0018720812442086

Mehler, B., Reimer, B., Coughlin, J. F., and Dusek, J. A. (2009). Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. *Transp. Res. Rec.* 2138, 6–12. doi: 10.3141/2138-02

Mehler, B., Reimer, B., and Dusek, J. A. (2011). *Working Paper 2011-3B MIT AgeLab Delayed Digit Recall Task* (n-back), Cambridge, MA.

Michael, L., Passmann, S., and Becker, R. (2012). Electrodermal lability as an indicator for subjective sleepiness during total sleep deprivation. *J. Sleep Res.* 21, 470–478. doi: 10.1111/j.1365-2869.2011.00984.x

Miró, E., Cano-Lozano, M. C., and Buela-Casal, G. (2002). Electrodermal activity during total sleep deprivation and its relationship with other activation and performance measures. *J. Sleep Res.* 11, 105–112. doi: 10.1046/j.1365-2869.2002.00286.x

Mitchell, D. J., McNaughton, N., Flanagan, D., and Kirk, I. J. (2008). Frontal-midline theta from the perspective of hippocampal "theta." *Prog. Neurobiol.* 86, 156–185. doi: 10.1016/j.pneurobio.2008.09.005

Mizuno, K., Tanaka, M., Yamaguti, K., Kajimoto, O., Kuratsune, H., and Watanabe, Y. (2011). Mental fatigue caused by prolonged cognitive load associated with sympathetic hyperactivity. *Behav. Brain Funct.* 7:17. doi: 10.1186/1744-9081-7-17

Murphey, Y., Kochhar, D. S., and Xie, Y. (2019). *Driver Workload in an Autonomous Vehicle. SAE Technical Papers, 2019-01-0872*. doi: 10.4271/2019-01-0872

Muttart, J. W., Fisher, D. L., Knodler, M., and Pollatsek, A. (2007). Driving without a clue: evaluation of driver simulator performance during hands-free cell phone operation in a work zone. *Transp. Res. Record.* 2018, 9–14. doi: 10.3141/2018-02

Nakano, T., Kato, M., Morito, Y., Itoi, S., and Kitazawa, S. (2013). Blink-related momentary activation of the default mode network while viewing videos. *Proc. Natl. Acad. Sci. U.S.A.* 110, 702–706. doi: 10.1073/pnas.1214804110

Niezgoda, M., Tarnowski, A., Kruszewski, M., and Kamiński, T. (2015). Towards testing auditory-vocal interfaces and detecting distraction while driving: a comparison of eye-movement measures in the assessment of cognitive workload. *Transp. Res. F Traffic Psychol. Behav.* 32, 23–34. doi: 10.1016/j.trf.2015.04.012

Nilsson, E. J., Aust, M. L., Engström, J., Svanberg, B., and Lindén, P. (2018). Effects of cognitive load on response time in an unexpected lead vehicle braking scenario and the detection response task (DRT). *Transp. Res. F Traffic Psychol. Behav.* 59, 463–474. doi: 10.1016/j.trf.2018.09.026

Nilsson, E. J., Victor, T., Ljung Aust, M., Svanberg, B., Lindén, P., and 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. *Transp. Res. F Traffic Psychol. Behav.* 75, 439–445. doi: 10.1016/j.trf.2020.09.013

Noh, Y., Kim, S., Jang, Y. J., and Yoon, Y. (2021). Modeling individual differences in driver workload inference using physiological data. *Int. J. Automot. Technol.* 22, 201–212. doi: 10.1007/s12239-021-0020-8

Nunes, L., and Recarte, M. A. (2002). Cognitive demands of hands-free-phone conversation while driving. *Transp. Res. F Traffic Psychol. Behav.* 5, 133–144. doi: 10.1016/S1369-8478(02)00012-8

O'Donnell, F. T., and Eggemeier, R. D. (1986). "Workload assessment methodology," in *Handbook of Perception and Human Performance*, 2nd Edn, eds K. R. Boff, L. Kaufman, and J. P. Thomas (New York, NY: John Wiley & Sons), 1–49.

Park, G., and Thayer, J. F. (2014). From the heart to the mind: cardiac vagal tone modulates top-down and bottom-up visual perception and attention to emotional stimuli. *Front. Psychol.* 5:278. doi: 10.3389/fpsyg.2014.00278

Paxion, J., Galy, E., and Berthelon, C. (2014). Mental workload and driving. *Front. Psychol.* 5:1344. doi: 10.3389/fpsyg.2014.01344

Pedrotti, M., Mirzaei, M. A., Tedesco, A., Chardonnet, J. R., Mérienne, F., Benedetto, S., et al. (2014). Automatic stress classification with pupil diameter analysis. *Int. J. Hum. Comput. Interact.* 30, 220–236. doi: 10.1080/10447318.2013.848320

Perrier, J., Jongen, S., Vuurman, E., Bocca, M. L., Ramaekers, J. G., and Vermeeren, A. (2016). Driving performance and EEG fluctuations during on-the-road driving following sleep deprivation. *Biol. Psychol.* 121, 1–11. doi: 10.1016/j.biopsycho.2016.09.010

Persson, A., Jonasson, H., Fredriksson, I., Wiklund, U., and Ahlstrom, C. (2020). Heart rate variability for classification of alert versus sleep deprived drivers in real road driving conditions. *IEEE Trans. Intell. Transp. Syst.* 22, 3316–3325. doi: 10.1109/TITS.2020.2981941

Peruzzini, M., Tonietti, M., and Iani, C. (2019). Transdisciplinary design approach based on driver's workload monitoring. *J. Indust. Inform. Integrat.* 15, 91–102. doi: 10.1016/j.jii.2019.04.001

Pfurtscheller, G., Stancák A. Jr., and Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band - an electrophysiological correlate of cortical idling: a review. *Int. J. Psychophysiol.* 24, 39–46. doi: 10.1016/S0167-8760(96)00066-9

Pion-Tonachini, L., Kreutz-Delgado, K., and Makeig, S. (2019). ICLabel: an automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage* 198, 181–197. doi: 10.1016/j.neuroimage.2019.05.026

Posada-Quintero, H. F., and Chon, K. H. (2020). Innovations in electrodermal activity data collection and signal processing: a systematic review. *Sensors* 20:479. doi: 10.3390/s20020479

Prabhakar, G., Mukhopadhyay, A., Murthy, L., Modiksha, M., Sachin, D., and Biswas, P. (2020). Cognitive load estimation using ocular parameters in automotive. *Transp. Eng.* 2:100008. doi: 10.1016/j.treng.2020.100008

Putze, F., Jarvis, J. P., and Schultz, T. (2010). "Multimodal recognition of cognitive workload for multitasking in the car," in *Proceedings - International Conference on Pattern Recognition* (Istanbul), 3748–3751. doi: 10.1109/ICPR.2010.913

Quintana, D. S., and Heathers, J. A. J. (2014). Considerations in the assessment of heart rate variability in biobehavioral research. *Front. Psychol.* 5:805. doi: 10.3389/fpsyg.2014.00805

Recarte, M. A., and Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *J. Exp. Psychol.* 6, 31–43. doi: 10.1037/1076-898X.6.1.31

Recarte, M. A., Pérez, E., Conchillo, Á., and Nunes, L. M. (2008). Mental workload and visual impairment: differences between pupil, blink, and subjective rating. *Span. J. Psychol.* 11, 374–385. doi: 10.1017/S1138741600004406

Reimer, B., and Mehler, B. (2011). The impact of cognitive workload on physiological arousal in young adult drivers: a field study and simulation validation. *Ergonomics* 54, 932–942. doi: 10.1080/00140139.2011.604431

Reimer, B., Mehler, B., Wang, Y., and Coughlin, J. F. (2012). A field study on the impact of variations in short-term memory demands on drivers' visual attention and driving performance across three age groups. *Hum. Factors* 54, 454–468. doi: 10.1177/0018720812437274

Rendon-Velez, E., van Leeuwen, P. M., Happee, R., Horváth, I., van der Vegte, W. F., and de Winter, J. C. F. (2016). The effects of time pressure on driver performance and physiological activity: a driving simulator study. *Transp. Res. F Traffic Psychol. Behav.* 41, 150–169. doi: 10.1016/j.trf.2016.06.013

Richter, M., and Slade, K. (2017). Interpretation of physiological indicators of motivation: caveats and recommendations. *Int. J. Psychophysiol.* 119, 4–10. doi: 10.1016/j.ijpsycho.2017.04.007

Rodrígue-Ibáñez, N., García-González, M. A., Fernández-Chimeno, M., and Ramos-Castro J. (2011). Drowsiness detection by thoracic effort signal analysis in real driving environments. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc*. 2011, 6055–6058. doi: 10.1109/IEMBS.2011.6091496

Sadaghiani, S., and Kleinschmidt, A. (2016). Brain networks and α-Oscillations: structural and functional foundations of cognitive control. *Trends Cogn. Sci.* 20, 805–817. doi: 10.1016/j.tics.2016.09.004

Samuels, E., and Szabadi, E. (2008). Functional neuroanatomy of the noradrenergic locus coeruleus: its roles in the regulation of arousal and autonomic function part i: principles of functional organisation. *Curr. Neuropharmacol.* 6, 235–253. doi: 10.2174/157015908785777229

Saxby, D. J., Matthews, G., Warm, J. S., Hitchcock, E. M., and Neubauer, C. (2013). Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. *J. Exp. Psychol.* 19, 287–300. doi: 10.1037/a0034386

Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., and Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accid. Anal. Prevent.* 41, 1087–1093. doi: 10.1016/j.aap.2009.06.007

Schneider, W., and Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. detection, search, and attention. *Psychol. Rev.* 84, 1–66. doi: 10.1037/0033-295X.84.1.1

Schoofs, D., Preuß, D., and Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology* 33, 643–653. doi: 10.1016/j.psyneuen.2008.02.004

Sescousse, G., Ligneul, R., van Holst, R. J., Janssen, L. K., de Boer, F., Janssen, M., et al. (2018). Spontaneous eye blink rate and dopamine synthesis capacity: preliminary evidence for an absence of positive correlation. *Europ. J. Neurosci.* 47, 1081–1086. doi: 10.1111/ejn.13895

Shaffer, F., and Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Front. Public Health* 5:258. doi: 10.3389/fpubh.2017.00258

Silvani, A., Calandra-buonaura, G., Dampney, R. A. L., and Cortelli, P. (2016). Brain – heart interactions : physiology and clinical implications. *Phil. Trans. R. Soc. A* 374:20150181. doi: 10.1098/rsta.2015.0181

Simon, M., Schmidt, E. A., Kincses, W. E., Fritzsche, M., Bruns, A., Aufmuth, C., et al. (2011). Eeg alpha spindle measures as indicators of driver fatigue under real traffic conditions. *Clin. Neurophysiol.* 122, 1168–1178. doi: 10.1016/j.clinph.2010.10.044

Sonnleitner, A., Simon, M., Kincses, W. E., Buchner, A., and Schrauf, M. (2012). Alpha spindles as neurophysiological correlates indicating attentional shift in a simulated driving task. *Int. J. Psychophysiol.* 83, 110–118. doi: 10.1016/j.ijpsycho.2011.10.013

Stephens, C. L., Christie, I. C., and Friedman, B. H. (2010). Autonomic specificity of basic emotions: Evidence from pattern classification and cluster analysis. *Biol. Psychol.* 84, 463–473. doi: 10.1016/j.biopsycho.2010.03.014

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

Strauss, M. E., and Smith, G. T. (2009). Construct validity: advances in theory and methodology. *Annu. Rev. Clin. Psychol.* 5, 1–25. doi: 10.1146/annurev.clinpsy.032408.153639

Strayer, D. L., and Fisher, D. L. (2016). SPIDER: a framework for understanding driver distraction. *Hum. Factors* 58, 5–12. doi: 10.1177/0018720815619074

Strayer, D. L., Martinez, M., Cooper, J. M., and Drews, F. A. (2006). "Brain waves suppressed by cell phone conversations," in *Proceedings of the Human Factors and Ergonomics Society* (San Francisco, CA), 2364–2367. doi: 10.1177/154193120605002203

Stuiver, A., Brookhuis, K. A., de Waard, D., and Mulder, B. (2014). Short-term cardiovascular measures for driver support: increasing sensitivity for detecting changes in mental workload. *Int. J. Psychophysiol.* 92, 35–41. doi: 10.1016/j.ijpsycho.2014.01.010

Szalma, J. L. (2008). "Individual differences in stress reaction," in *Performance Under Stress*, eds P. A. Hancock and J. L. Szalma (Hampshire: Ashgate), 323–357.

Takeda, Y., Sato, T., Kimura, K., Komine, H., Akamatsu, M., and Sato, J. (2016). Electrophysiological evaluation of attention in drivers and passengers: toward an understanding of drivers' attentional state in autonomous vehicles. *Transp. Res. F Traffic Psychol. Behav.* 42, 140–150. doi: 10.1016/j.trf.2016.07.008

Tanaka, M., Mizuno, K., Tajima, S., Sasabe, T., and Watanabe, Y. (2009). Central nervous system fatigue alters autonomic nerve activity. *Life Sci.* 84, 235–239. doi: 10.1016/j.lfs.2008.12.004

Tao, D., Tan, H., Wang, H., Zhang, X., Qu, X., and Zhang, T. (2019). A systematic review of physiological measures of mental workload. *Int. J. Environ. Res. Public Health* 16:2716. doi: 10.3390/ijerph16152716

Tatum, W. O. IV. (2014). *Handbook of EEG interpretation,* (2nd ed.), ed W. O. Tatum, IV. New York, NY: Springer Publishing Company. doi: 10.1891/9781617051807

Tejero, P., and Roca, J. (2021). Messages beyond the phone: processing variable message signs while attending hands-free phone calls. *Accid. Anal. Prevent.* 150:105870. doi: 10.1016/j.aap.2020.105870

Törnros, J., and Bolling, A. (2006). Mobile phone use - effects of conversation on mental workload and driving speed in rural and urban environments. *Transp. Res. F Traffic Psychol. Behav.* 9, 298–306. doi: 10.1016/j.trf.2006.01.008

Tortora, G. J., and Derrickson, B. (2007). *Principles of Anatomy and Physiology (11th ed.)*. Hoboken, NJ: John Wiley & Sons, Inc.

Tran, Y., Craig, A., Craig, R., Chai, R., and Nguyen, H. (2020). The influence of mental fatigue on brain activity: evidence from a systematic review with meta-analyses. *Psychophysiology* 57:e13554. doi: 10.1111/psyp.13554

Van Acker, B. B., Parmentier, D. D., Vlerick, P., and Saldien, J. (2018). Understanding mental workload: from a clarifying concept analysis toward an implementable framework. *Cogn. Technol. Work* 20, 351–365. doi: 10.1007/s10111-018-0481-3

van der Wel, P., and van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: a review. *Psychon. Bull. Rev.* 25, 2005–2015. doi: 10.3758/s13423-018-1432-y

Victor, T., Dozza, M., Bärgman, J., Boda, C.-N., Engström, J., Flannagan, C., et al. (2015). *Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk. Report S2-S08A-RW-1*. (Washington, DC). doi: 10.17226/22297

Vidaurre, C., Sander, T. H., and Schlögl, A. (2011). BioSig: the free and open source software library for biomedical signal processing. *Comput. Intell. Neurosci.* 2011:935364. doi: 10.1155/2011/935364

Visnovcova, Z., Mestanik, M., Gala, M., Mestanikova, A., and Tonhajzerova, I. (2016). The complexity of electrodermal activity is altered in mental cognitive stressors. *Comput. Biol. Med.* 79, 123–129. doi: 10.1016/j.compbiomed.2016.10.014

Vlemincx, E., Abelson, J. L., Lehrer, P. M., Davenport, P. W., Van Diest, I., and Van Den Bergh, O. (2013). Respiratory variability and sighing: a psychophysiological reset model. *Biol. Psychol.* 93, 24–32. doi: 10.1016/j.biopsycho.2012.12.001

Vlemincx, E., Van Diest, I., and Van den Bergh, O. (2012). A sigh following sustained attention and mental stress: effects on respiratory variability. *Physiol. Behav.* 107, 1–6. doi: 10.1016/j.physbeh.2012.05.013

Wang, Y. K., Jung, T. P., and Lin, C. T. (2018). Theta and alpha oscillations in attentional interaction during distracted driving. *Front. Behav. Neurosci.* 12:3. doi: 10.3389/fnbeh.2018.00003

Wascher, E., Arnau, S., Gutberlet, I., Karthaus, M., and Getzmann, S. (2018). Evaluating pro- and re-active driving behavior by means of the EEG. *Front. Hum. Neurosci.* 12:205. doi: 10.3389/fnhum.2018.00205

Wiberg, H., Nilsson, E., Lindén, P., Svanberg, B., and Poom, L. (2015). Physiological responses related to moderate mental load during car driving in field conditions. *Biol. Psychol.* 108, 115–125. doi: 10.1016/j.biopsycho.2015.03.017

Wijayaratna, K. P., Cunningham, M. L., Regan, M. A., Jian, S., Chand, S., and Dixit, V. V. (2019). Mobile phone conversation distraction: understanding differences in impact between simulator and naturalistic driving studies. *Accid. Anal. Prevent.* 129, 108–118. doi: 10.1016/j.aap.2019.04.017

Winn, B., Whitaker, D., Elliott, D. B., and Phillips, N. J. (1994). Factors affecting light-adapted pupil size in normal human subjects. *Invest. Ophthalmol. Visual Sci.* 35, 1132–1137.

Xie, L., Duan, M., and Li, W. (2020). "Discriminant model of driving distraction during mobile phone conversation based on eye movements," in *Green, Smart and Connected Transportation Systems: Lecture Notes in Electrical Engineering*, Vol. 617, eds W. Wang, M. Baumann, and X. Jiang (Singapore: Springer), 721–731. doi: 10.1007/978-981-15-0644-4_56

Young, M. S., Brookhuis, K. A., Wickens, C. D., and Hancock, P. A. (2015). State of science: mental workload in ergonomics. *Ergonomics* 58, 1–17. doi: 10.1080/00140139.2014.956151

Zénon, A. (2019). Eye pupil signals information gain. *Proc. Biol. Sci.* 286:20191593. doi: 10.1098/rspb.2019.1593

Zokaei, M., Jafari, M. J., Khosrowabadi, R., Nahvi, A., Khodakarim, S., and Pouyakian, M. (2020). Tracing the physiological response and behavioral performance of drivers at different levels of mental workload using driving simulators. *J. Safety Res.* 72, 213–223. doi: 10.1016/j.jsr.2019.12.022

Conflict of Interest: EN, BS, and ML are employed by Volvo Car Corporation.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroergonomics

ORIGINAL RESEARCH
published: 24 February 2022
doi: 10.3389/fnrgo.2022.836518

Investigating Differences in Behavior and Brain in Human-Human and Human-Autonomous Vehicle Interactions in Time-Critical Situations

*Anirudh Unni[1], Alexander Trende[2], Claire Pauley[1], Lars Weber[2], Bianca Biebl[3], Severin Kacianka[4], Andreas Lüdtke[2], Klaus Bengler[3], Alexander Pretschner[4], Martin Fränzle[2,5] and Jochem W. Rieger[1]**

[1] Department of Psychology, University of Oldenburg, Oldenburg, Germany, [2] OFFIS Institute for Information Technology, Division of Transportation Research, Oldenburg, Germany, [3] Chair of Ergonomics, Technical University of Munich, Garching, Germany, [4] Chair of Software and Systems Engineering, Technical University of Munich, Garching, Germany, [5] Department of Computer Science, University of Oldenburg, Oldenburg, Germany

OPEN ACCESS

Edited by:
*Giovanni Vecchiato,
National Research Council of Italy, Italy*

Reviewed by:
*Tuomo Kujala,
University of Jyväskylä, Finland
Carryl L. Baldwin,
Wichita State University, United States*

***Correspondence:**
*Jochem W. Rieger
jochem.rieger@uni-oldenburg.de*

Specialty section:
This article was submitted to Cognitive Neuroergonomics, a section of the journal Frontiers in Neuroergonomics

Received: *15 December 2021*
Accepted: *01 February 2022*
Published: *24 February 2022*

Citation:
Unni A, Trende A, Pauley C, Weber L, Biebl B, Kacianka S, Lüdtke A, Bengler K, Pretschner A, Fränzle M and Rieger JW (2022) Investigating Differences in Behavior and Brain in Human-Human and Human-Autonomous Vehicle Interactions in Time-Critical Situations. Front. Neuroergon. 3:836518. doi: 10.3389/fnrgo.2022.836518

Some studies provide evidence that humans could actively exploit the alleged technological advantages of autonomous vehicles (AVs). This implies that humans may tend to interact differently with AVs as compared to human driven vehicles (HVs) with the knowledge that AVs are programmed to be risk-averse. Hence, it is important to investigate how humans interact with AVs in complex traffic situations. Here, we investigated whether participants would value interactions with AVs differently compared to HVs, and if these differences can be characterized on the behavioral and brain-level. We presented participants with a cover story while recording whole-head brain activity using fNIRS that they were driving under time pressure through urban traffic in the presence of other HVs and AVs. Moreover, the AVs were programmed defensively to avoid collisions and had faster braking reaction times than HVs. Participants would receive a monetary reward if they managed to finish the driving block within a given time-limit without risky driving maneuvers. During the drive, participants were repeatedly confronted with left-lane turning situations at unsignalized intersections. They had to stop and find a gap to turn in front of an oncoming stream of vehicles consisting of HVs and AVs. While the behavioral results did not show any significant difference between the safety margin used during the turning maneuvers with respect to AVs or HVs, participants tended to be more certain in their decision-making process while turning in front of AVs as reflected by the smaller variance in the gap size acceptance as compared to HVs. Importantly, using a multivariate logistic regression approach, we were able to predict whether the participants decided to turn in front of HVs or AVs from whole-head fNIRS in the decision-making phase for every participant (mean accuracy = 67.2%, SD = 5%). Channel-wise univariate fNIRS analysis revealed increased brain activation differences for turning in front of AVs compared to HVs in brain areas that represent the valuation of actions taken during decision-making. The insights provided here may be useful for the development of control systems to assess interactions in future mixed traffic environments involving AVs and HVs.

Keywords: human-autonomous vehicle interaction, whole-head fNIRS, multivariate logistic ridge regression, valuation of actions, decision-making

INTRODUCTION

A majority of vehicle accidents are caused by human errors (Singh, 2018). A long-held belief is that the introduction of autonomous vehicles (AVs) in driving will reduce human errors, leading to an overall improvement in terms of driving performance and safety for all traffic participants. However, until a time comes when only AVs travel on roads, human driven vehicles (HVs) and AVs will co-exist in traffic environments. In such mixed traffic environments, the interaction between humans and autonomous agents remains extremely important. This is of concern regarding not only the humans using AVs, but also regarding the interaction between HVs and AVs.

A key aspect for a safe and seamless interaction between HVs and AVs is how human's actions are influenced by AVs in mixed traffic environments. In fact, some studies have shown that pedestrians and human drivers could actively exploit the alleged technological advantages of AVs. For example, the pedestrian or the human driver knows that AVs are programmed to be risk-averse and stop immediately if it detects an obstacle in its path. Armed with this knowledge, drivers and pedestrians may act with impunity while interacting with AVs. Several studies have reported a shift in behavior when humans are interacting with autonomous agents compared to other human agents suggesting that humans might evaluate their own actions differently depending on the type of traffic agent involved. For example, Trende et al. (2019) showed that in time-critical situations, drivers had a significantly higher gap acceptance probability for turning in front of an AV as compared to HV. Moreover, Millard-Ball (2018) showed that pedestrians took advantage of a mildly-mannered AV knowing that the AV will yield at crosswalks, and they can hence cross the road with impunity. Similar results were reported by Liu et al. (2020) where drivers revealed greater intentions to drive aggressively while interacting with AVs as compared to HVs. Such actions of the driver could be constituted as *"misuse of automation,"* a term coined by Parasuraman and Riley (1997). One such type of automation misuse potentially leading to dangerous situations when interacting with AVs is an *"overreliance"* on the automation system (Parasuraman and Manzey, 2010). Overreliance occurs when a driver tends to rely uncritically on the automation without recognizing its limitations or fails to monitor the automation system's behavior (Saffarian et al., 2012; Cunningham and Regan, 2015).

The assessment of safety-critical situations in complex traffic requires significant cognitive resources to form a mental representation of the situation, to identify potentially critical interaction partners and to predict their behavior. The correct estimation and expectation of other's behavior plays a crucial role for safe interaction. In situations where the HV and AV need to interact directly, the driver may tend to underestimate the reaction time of an AV leading to a risky maneuver. The prediction of the AV's behavior in complex traffic situations is based on the driver's mental model of the AV. Mental models are internal representations of a system concerning its characteristics, potentials and limitations that are mainly formed by interacting with the system (Kurpiers et al., 2020). Such mental models can influence information processing, valuation of actions and the resulting decision to act in human-autonomous vehicle interactions. However, it is hard to evaluate such mental models due to their implicit nature and more objective measures are required.

Neurophysiological measurements allow for an objective tracking of cognitive processes such as decision-making. Spatially resolved brain activation measures can be more specific to decision-making processes as they are recorded at the location where these cognitive processes are manifested. This allows us to unravel what goes on in a driver's brain while performing decision-making interactions with technical systems such as AV. Until now, a solid number of neuroimaging studies have been conducted that revealed human brain areas involved in decision-making and characterized their responses in game theoretic frameworks. Much progress has been made in defining game-theoretic building blocks of human decision-making models and implementing these blocks in executable cognitive architectures such as ACT-R (Taatgen et al., 2005). Moreover, neurophysiological research has revealed neural correlates for action-based value signals for reward related decision-making tasks. Some of these brain areas include the prefrontal cortices such as the dorsolateral prefrontal cortex (dlPFC), ventromedial prefrontal cortex (vmPFC), frontal cingulate, anterior orbito- and mediofrontal cortices (Sanfey, 2007; Lee, 2008; Rangel et al., 2008; Gläscher et al., 2009; Ruff and Fehr, 2014). However, very few studies in this field have actually attempted to predict human decision-making interactions from brain activation in realistic situations. Hollmann et al. (2011) employed real-time functional MRI to predict online decisions during social interactions in the ultimatum game from brain activation and to reveal brain areas that signal whether offers were subjectively perceived as unfair. These approaches have been extended from relatively simple operant conditioning in laboratory environments (Schultz, 2002) to decision-making in social context (Sanfey et al., 2003; Sanfey, 2007). However, to the best of our knowledge, no neurophysiological study has compared how interactions with other humans or technical systems such as AVs are reflected in characterizing neural correlates for decision-making in realistic scenarios such as driving using fNIRS.

In this study, we use turning at an intersection as a safety-critical traffic situation, where the driver must directly interact with other traffic participants. Previous studies have reported that between 30 and 40% of crashes are located at or near intersections even though these situations represent only a small percentage of the entire road infrastructure (Tay and Rifaat, 2007; Choi, 2010; Gerstenberger, 2015). A lacking consideration for other road users is the primary reason for accidents when turning according to a report by the German Federal Highway Research Institute (BASt) on intersection-related crash factors (Vollrath et al., 2006; Biebl and Bengler, 2021). When a vehicle stops at an intersection, the driver must observe the oncoming traffic stream before accepting a gap and turning into the desired lane. The gap acceptance problem is one of the main causes for stop-controlled intersection accidents (Yan et al., 2007). Several studies have conducted field observations or driving simulator studies to investigate gap acceptance in these situations (Ragland et al., 2006; Yan et al., 2007; Lord-Attivor and Jha, 2012), leading some of them to predict gap acceptance using statistical models. Lord-Attivor and Jha (2012) collected data from Nigerian intersections and proposed a binary logit model to model gap acceptance behavior. Furthermore, Ragland et al. (2006) analyzed video recordings of five intersections to determine gap acceptance statistics and proposed a logit model predicting gap acceptance probability. Such models can help to design and develop driving assistance decision support systems, which can potentially reduce the number of traffic accidents at the intersections.

The objective of this study is to investigate if there is a difference between a human driver's valuation of actions when an interaction involves technical systems such as AVs as compared to similar interactions with other human beings. In a second step, this paper aims to examine whether these potential differences in human-human and human-autonomous vehicle interactions can be characterized from behavior and neurophysiological whole-head fNIRS brain activation measurements. For this purpose, we conducted an fNIRS-driving simulator study. We measured whole-head brain activation using high density fNIRS throughout the entire driving time to identify neural correlates associated with the valuation of actions during decision-making in the turning situations in human-human and human-autonomous vehicle interactions. We presented the participants with a cover story that they were driving under time pressure through urban traffic in the presence of other HVs and AVs, that the AVs were programmed defensively to avoid collisions and that they had faster braking reaction times than HVs. Participants would receive a monetary reward if they managed to finish the driving block by avoiding risky driving maneuvers within a given time limit. The participants were repeatedly confronted with a left-lane turning situation at unsignalized intersections where they had to decide to turn in front of a HV or an AV. We hypothesize that under time pressure, there is more considerate behavior while interacting with HVs than with AVs as for the latter, there is no safety-critical consequence of one's own actions due to the driver's expectation that AVs drive more cautiously making them more predictable in their driving behavior as compared to HVs. This would be reflected in reduced safety margins (e.g., gap sizes) and increased certainty during the decision-making process while interacting with AVs as compared to HVs. Based on previous research (Sanfey, 2007; Rangel et al., 2008; Ruff and Fehr, 2014), we hypothesize that human-autonomous vehicle interactions cause increased activation modulations in the prefrontal areas such as the dorsolateral and ventrolateral prefrontal, ventromedial prefrontal, frontal midline brain areas and the anterior cingulate cortex, since these brain areas are thought to represent the consideration of values of actions taken during decision-making.

MATERIALS AND METHODS

Participants

Thirteen volunteers (7 females) aged 21–29 years (Mean $\pm$ SD $= 23.8 \pm 2.61$) participated in the study. The participants had a mean driving experience of 5.8 years (SD $= 2.5$). All participants possessed a valid German driving license and gave written informed consent to participate prior to the experiment in accordance with the Declaration of Helsinki. The Ethics Committee of the Carl von Ossietzky University, Oldenburg approved the experimental procedure. Participants received a financial reimbursement of 10 € per hour.

Experimental Set-Up

The experiment was performed in a full-scale fixed-base driving simulator, which offered a 150° field of view (**Figure 1**). The driving simulator contained a realistic vehicle mock-up. The driving simulator software SILAB (Krueger et al., 2005) was used to simulate the driving scenario. The participants controlled the mock-up car in the driving simulation via a standard interface consisting of a throttle, brake pedal and steering wheel. Behavioral data, such as acceleration, velocity and steering wheel angle were recorded via SILAB.

Participants' brain activation was measured using a high density, whole-head fNIRS system throughout the entire driving time. fNIRS uses the principle of neurovascular coupling where the neuronal activity is linked to related absorption changes in the sub-surface tissues in localized cerebral blood flow

FIGURE 1 | Virtual reality lab driving simulator at OFFIS Institute of Information Technology, Oldenburg—photograph of experimental setup. The participant's brain activity is measured with whole-head fNIRS system while they are driving in the urban traffic.

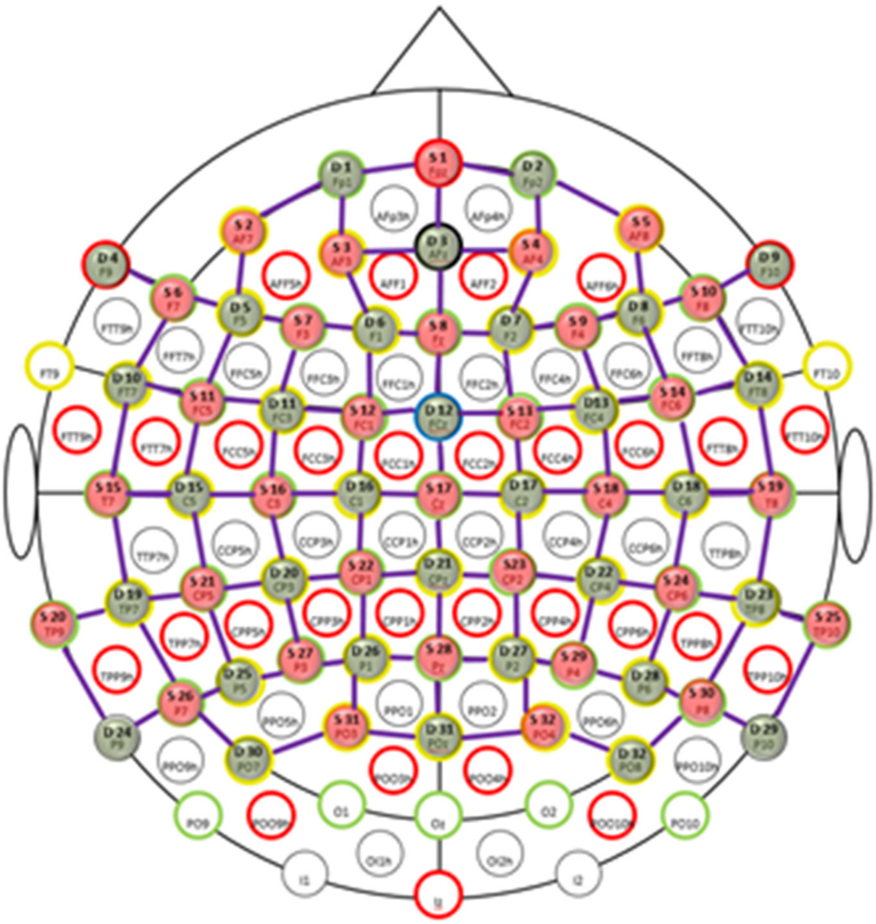


FIGURE 2 | FNIRS probe placement. Topologic layout of the emitters (red disks), detectors (green disks) and the fNIRS channels (purple lines) on a standard 10–20 EEG system. Figure reproduced from NIRStar 15.0 data acquisition software with permission from NIRx Medical Technologies, USA.

by measuring local concentration changes of oxyhaemoglobin (HbO) and deoxyhaemoglobin (HbR) as correlates of functional brain activity using the modified Beer-Lambert law (Villringer et al., 1993; Sassaroli and Fantini, 2004). We used the NIRScout Extended system (NIRx Medical Technologies) to acquire fNIRS data. The system uses two wavelengths of 760 nm and 850 nm and outputs relative concentration changes of HbO and HbR. Thirty-two optical emitters and detectors were used to obtain close to whole-head coverage. In total, 107 channels (combinations of emitters and receivers) were used to acquire fNIRS data at a sampling frequency of 1.955 Hz (**Figure 2**). The average distance between an emitter and detector was ∼3.5 cm.

Both the fNIRS and driving simulator data were trigger-synchronized during the driving task.

Experimental Design

The driving simulation featured multiple left-turn maneuvers with oncoming traffic in an urban environment. The oncoming vehicles drove at a speed of 50.4 km/h (equivalent to 14 m/s), the speed limit for most urban roads in Germany. Due to a STOP sign, the subject vehicle had to stop at the intersection before accepting a gap and turning into the desired lane (**Figure 3**). A "gap" represents the opportunity to turn left before an oncoming vehicle. Every gap has an associated gap size that represents the time in seconds that passes after the first of two successive, oncoming vehicles passes the intersection until the second vehicle passed the intersection. The driver faces a series of gaps of different sizes while waiting at the intersection and has the choice to either accept or reject a given gap. Accepting a gap means that the driver completes a left-turn maneuver.

The lane of oncoming traffic was bent slightly to the right (**Figure 3**). This makes the estimation of the gap sizes between oncoming vehicles easier. The simulated traffic consisted of human driven vehicles (HVs) and autonomous vehicles (AVs) (Trende et al., 2019). The AVs were always yellow cars without a virtual human model visible inside the car. The HVs were represented by cars of other colors (except yellow) where the virtual human model was clearly visible. Before the experiment, the participants were instructed how the AVs will look in the simulation. The participants were told that the AVs are

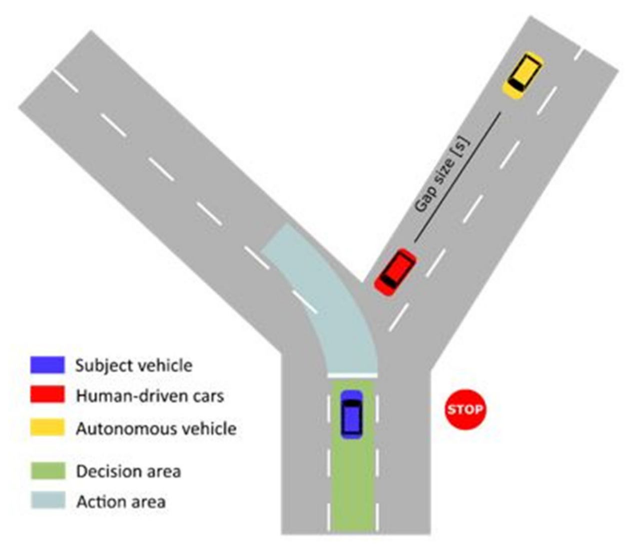


FIGURE 3 | Sketch of intersection. The gap size between two oncoming vehicles defined as the time that passes after the first oncoming vehicle has crossed the intersection until the second vehicle has crossed the intersection.

programmed to use a defensive, risk-avoiding driving style (Millard-Ball, 2018). However, in reality, both HV and AV followed the same driving behavior. The number of AVs in the simulation was lower than the number of HVs since automated driving is a novel technique and only a few AVs are available on the market. AVs represented 15% of the simulated cars. While waiting at the intersection in front of the STOP sign, between eight and 10 cars approached the intersection.

We followed Ragland et al. (2006) to design realistic traffic situations. The authors used video data from five intersections in the USA to find the distributions of gaps smaller than 12 s between subsequent cars at intersections. They found that most gaps were 4 s or shorter with the most frequent gap being 2 s. Overall, the gap size distribution could be modeled as a lognormal distribution. We designed the distribution of gaps in our study according to these findings. We decided to present gap sizes between 1 and 6 s during the experiment. As suggested in Yan et al. (2007), we designed the traffic in such a way that the first oncoming vehicles have lower gap sizes. This helps to find minimal gap size acceptances for participants and assures that a suitable gap size for each participant's preferences was presented. We split the oncoming traffic in two groups: The first 4–6 cars have a gap size from a range of 1–3 s. A larger gap in the range of 3.5–6 s was placed after fifth to tenth car, respectively. No vehicles appeared after the tenth car. The sequence of the type of involved traffic agent (i.e., HV or AV) among the stream of oncoming traffic encountered at the intersection varied during one experiment but remained the same for all participants. We performed a training session before the experiment in which the participants drove a short scenario consisting of rural roads and 11 intersections, which took around 10 min. The purpose of the training scenario was to get accustomed to the virtual reality environment and simulator dynamics.

In the experimental session, participants drove 100 intersections consisting of 10 driving blocks with 10 intersections per block. The whole session lasted around 70 min. The participants were asked to stop after 10 intersections and had a break of 1–1.5 min. Time pressure was applied during each block of the experiment. If the participants managed to reach the end of the 10^{th} intersection in a block within 5:30 min, they received a bonus of 1€ per block. The timer and intersection counter were displayed as a Heads-up-Display (HUD) in the simulation. To reach the end of the scenario within the block in the given time limit, the participants had to take at least some of the gaps while waiting at intersections. In principle, participants could have waited until the end of the oncoming stream of vehicles before deciding to turn. However, the time constraints introduced by the bonus discouraged participants to employ such a strategy. Across all participants, only 2 out of the 1,200 turning maneuvers were performed after the last car when the oncoming traffic had already passed.

After the experimental session, the participants were asked to fill out a questionnaire with 4 qualitative questions about trust in AV. The participants choose a score between 1 and 6 for the quantitative items. They rated 4 items related to trust in AV: "I accept AVs on the roads"; "AVs are safer than HVs"; "I trust AVs more than HVs" and "I behaved differently in my interactions with AVs compared to HVs."

Data Analysis

The data analysis section consists of three parts: analysis of the driving behavior, analysis of the neurophysiological data and analysis of the questionnaire.

Behavior Parameters

As presented in other studies (Fitzpatrick, 1991; Ragland et al., 2006), we calculate the gap acceptance probability for each gap size over all participants. The gap acceptance data was extracted based on the positional data of the subject vehicle and oncoming vehicles. We fitted a logistic model to the gap acceptance probability of the participants. The gap acceptance probability was calculated for gap sizes in 0.35 s steps. The logistic model had the following form and two regression parameters to fit.

$$P\left(X,\ m,\ w\right) = \frac{1}{\left(1 + exp\left(-2\log\left(\frac{1}{0.05} - 1\right)\frac{X-m}{w}\right)\right)} \quad (1)$$

Here, "X" represents the gap size, "m" is the threshold indicating a 50% gap acceptance and "w" is the width describing the difference between 5 and 95% point of the model. *MATLAB 2020* and the *psignifit 4* toolbox (Schütt et al., 2015) were used for fitting the logistic model to the data.

fNIRS Data Pre-processing

The raw fNIRS data are influenced not only by cortical brain activity but also by other systemic physiological artifacts (cardiac artifacts, respiration rate, and Mayer waves) and movement artifacts causing the signal to be noisy. We pre-processed the raw fNIRS data using the nirsLAB analysis package to reduce the influence of these artifacts (Xu et al., 2014). First, a "*coefficient of variation*" (CV) was computed which is a measure for the signal-to-noise ratio (SNR) from the unfiltered raw data. The CV was calculated as the ratio between the standard deviation and the mean of each NIRS channel over the entire duration of the experiment (Schmitz et al., 2005; Schneider et al., 2011). All channels with a CV >20% were excluded from further analysis. Moreover, we performed band pass filtering of the raw fNIRS data with a high cut-off frequency at 0.1 Hz to attenuate the effects of the above-mentioned physiological artifacts and instrument noise and a low cut-off frequency of 0.01 Hz to reduce the effects of very low frequency and gradual drift in the fNIRS data. Additionally, we visually inspected all channels and deleted those, which were excessively noisy with various spikes. Using these methods, on average, 99 fNIRS channels per participant were included in the subsequent analysis (SD = 8.7). Further, the modified Beer-Lambert's law was applied to convert the raw data from voltage (μV) to relative concentration change (*mmol/l*) (Sassaroli and Fantini, 2004).

The following fNIRS analysis was based on HbR signal, as HbR signals are considered to be less influenced by systemic physiological artifacts like cardiac pulsation, respiration, or Mayer wave fluctuations than HbO (Obrig et al., 2000; Zhang et al., 2005, 2009; Huppert et al., 2009; Suzuki, 2017). Moreover, other studies reported that HbR tends to correlate stronger with the blood oxygenated level dependent (BOLD) response than HbO (MacIntosh et al., 2003; Huppert et al., 2006; Schroeter et al., 2006; Foy et al., 2016).

We performed two types of analyses in order to better understand the neurophysiological activation differences as an index for differences in decision-making while turning in front of HV or AV and to characterize the contribution of these differences on a functional brain-level. The first type was a multivariate decoding modeling framework where our goal was to decode from the whole-head fNIRS activity whether the participant currently decided to turn in front of an HV or AV. The decision-making phase was defined as the decision to turn either in front of a HV or an AV along with the action to execute the decision. This phase corresponded to the timing 2 s before pressing the accelerator and initiating the decision to turn up to 2 s after beginning the turning maneuver for each trial. We selected this 4 s interval for the decision-making phase to account for the hemodynamic delay in the BOLD response measured by fNIRS. In the second type of analysis, we investigated the contribution of the brain activation features to such a decoding model that predicts human-human (turning in front of HVs) or human-autonomous (turning in front of AVs) interactions in the decision-making phase in a group-level by reporting the effect sizes for each fNIRS channel. The following sections provide further details about the methods to implement these analyses.

Multivariate Cross-Validated Prediction of Turning in Front of HV or AV

The goal of this analysis was to predict whether the participant decided to turn in front of a HV or AV from the pre-processed z-score normalized fNIRS data. First, since there were always more HV trials than AV trials, we balanced the trials by randomly selecting a sample of HV trials matching the number of AV trials available for each participant. Each timepoint (sampling frequency 1.955 Hz) in the 4 s time window during the decision-making phase while turning in front of a HV or an AV was considered as a single sample for the following classification.

The normalized fNIRS data was separated into train and test data. We calculated a multivariate binary logistic ridge regression model implemented in the Glmnet toolbox (Qian et al., 2013) within a 5-fold nested cross-validation on the samples to predict whether a particular timepoint in the fNIRS test data corresponded to human-human or human-autonomous interaction. The optimization of the hyperparameters (number of principal components (PCs) and regularization parameter λ) of the model was carried out in the training phase of the inner cross-validation loop. The outer cross-validation loop tested the generalization the logistic ridge regression model with the optimized hyperparameters to new data. This approach avoids overfitting of the model to the data and provides an estimate of how well the chosen decoding model would predict data that has not been seen previously by the model; for instance, in an online analysis (Hastie et al., 2009; Reichert et al., 2014).

The λ parameter, which determines the overall intensity of regularization of the logistic ridge regression model, was optimized by Glmnet using the training data within the cross-validation (Qian et al., 2013). We first performed a principal component analysis (PCA) on the training set. In this way, the fNIRS training data was transformed into a set of linearly uncorrelated variables called principal components (PCs). By this method, the first PC accounted for the largest variance in the data, and each successive component had the largest possible variance while maintaining orthogonality to the preceding components. The first PC has been shown to be linked to motion artifacts (Brigadoi et al., 2014), and was removed from further analysis. To increase the signal-to-noise ratio (SNR) and limit further analyses to the data explaining the most possible variance, all PCs with eigen values <0.7 were removed as recommended by Jolliffe (1972) on the Kaiser's rule (Kaiser, 1958). This resulted in an average of 13 PCs (SD = 2.4) per participant. The PCA eigen vectors of the training set was used to transform the test dataset in PC space.

Since the output of logistic regression can be interpreted as a class probability, all samples with a model output of $p \geq 0.5$ were assigned to the class "AV." This allowed us to calculate the rates at which the model correctly classified the two conditions. In this study, we report model accuracy, which indicates the proportion of correctly classified samples as either turning in front of an AV or a HV. The model accuracy was calculated as follows:

$$Accuracy\,(\%) = \frac{TP_{AV} + TP_{HV}}{TP_{AV} + TP_{HV} + FP_{AV} + FP_{HV}} * 100 \quad (2)$$

Here, TP refers to the number of true positives (number of samples correctly classified) and FP refers to the false positives (number of samples incorrectly classified) for the two conditions AV and HV. Further, we also calculated the F1-score, which is a combined harmonic average of the precision and recall measures of the model. The F1-score for AV condition was calculated as follows:

$$F1\text{-}score = \frac{2*TP_{AV}}{2*TP_{AV} + FP_{AV} + FP_{HV}} \quad (3)$$

The F1-score for HV condition was also calculated accordingly. We report the final mean model accuracy and F1-score for all participants.

Characterization of Brain Areas Predictive to Decision-Making Phase in Human-Human and Human-Autonomous Interactions

We aimed to characterize the separability of human-human or human-autonomous vehicle interactions from the channel-wise brain activation features used in the above described multivariate logistic ridge regression model. For this, we performed a channel-wise paired *t-test* from the preprocessed fNIRS data for the two conditions AV and HV on a single-subject level. To generalize the individual t-statistics brain maps to our test sample, the channel-wise single-subject *t-statistics* (t) were weighted with the participant's average model accuracy from the multivariate logistic ridge regression ($Accuracy_{mvr}$) to compute a weighted average t-statistics (t_{avg}) across the test sample (Unni et al., 2017) as shown below.

$$t_{avg}(i) = \frac{\sum_{i,\,n=1}^{i,\,n} t(i) * Accuracy_{mvr}(n)}{\sum_{1}^{n} Accuracy_{mvr}(n)} \quad (4)$$

Here, i refers to the total number of fNIRS channels and n indicates the total number of participants. We calculated *Cohen's d* for each channel from t_{avg} to indicate the effect sizes in sensor space.

$$Cohen's\,d(i) = \frac{(t_{avg}(i))^2}{\sqrt{df}} \quad (5)$$

Here, df refers to the degrees of freedom. We report these *Cohen's d* brain maps for the group-level analyses.

RESULTS

Data from one participant was excluded due to simulator sickness during the experiment. Thus, data from 12 participants are reported in the following sections.

Questionnaire Results

The results of the trust-related items from the questionnaire are shown in **Table 1**. The mean score for the overall trust-related items in the questionnaire was 3.8 (out of 6) indicating a high trust in AVs. The Cronbach alpha for the trust-related items was 0.78, indicating that these items have an acceptable reliability or internal consistency. Furthermore, 7 out of 12 participants stated that they turned in front of an AV preferably, when asked about a specific strategy for AVs. It is possible that these participants may feel that the interaction with a programmed vehicle is more controllable than with a human driver due to the AVs' perceived predictability. Another possible explanation could be that the participants potentially tried to exploit the defensive programming behavior and driving performance of AVs solely based on the cover story to gain a temporal advantage during the drive and achieve the bonus.

TABLE 1 | Results from the trust-related items of the post-experiment questionnaire.

Item	Mean score ± SD
I accept AVs on the roads.	4.2 ± 1.3
AVs are safer than HVs.	4.0 ± 1.5
I trust AVs more than HVs.	3.9 ± 1.4
I behaved differently in my interactions with AVs compared to HVs.	3.2 ± 1.2
Overall	3.8 ± 0.9

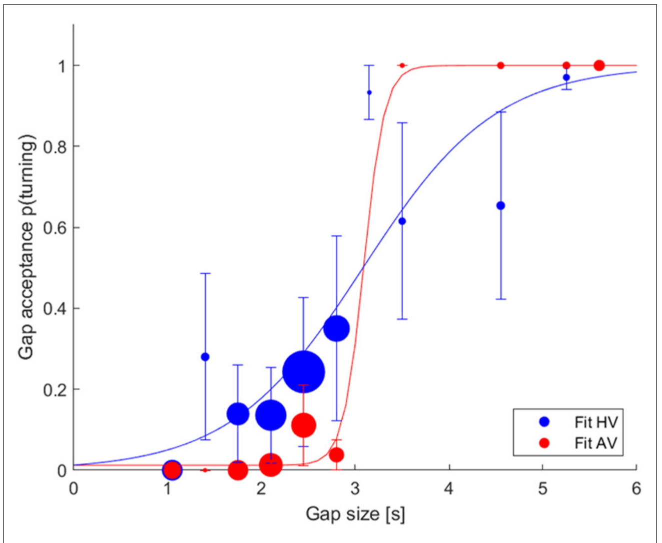


FIGURE 4 | Gap acceptance vs. gap size over all participants for HV (blue) and AV (red). The marker size represents the total number of events for the corresponding gap size. A logistic model was fit to each condition.

Behavioral Results

Figure 4 shows the gap acceptance probability in relation to the gap sizes by fitting the logistic model for HVs and AVs. Gap sizes were grouped in 0.35 second steps. The gap acceptance events were pooled over all participants. The models' widths (w) describing the difference between the 5 and 95% point of the model for AVs and HVs were $w_{AV} = 0.65$ s (0.59–1.25 s) and $w_{HV} = 4.17$ s (2.87–5.04 s), respectively. This is indicated by a steeper slope for AV as compared to HV in **Figure 4**.

The models' threshold values indicating the 50% gap acceptance for AVs and HVs were $m_{AV} = 3.09$ s (2.96–3.20 s) and

$m_{HV} = 3.08$ s (2.75–3.30 s), respectively. The difference in the threshold values is 0.04 m. The models' threshold values and their corresponding confidence intervals show an overlap, suggesting that these distributions do not differ significantly.

Prediction of Human-Human or Human-Autonomous Interaction From the Decoding Model

Using the multivariate logistic ridge regression model, we were able to predict the type of traffic agent (AV or HV) in the decision-making phase from whole-head fNIRS brain activation measurements with an average prediction accuracy and F1-score of 67.2% (SD = 3%) and 0.67 (SD = 0.05) respectively, across all participants. Prediction accuracies obtained in line with the measured dataset exceeded the 95% confidence interval (CI) for guessing for all participants. **Table 2** reports the individual prediction results for all participants along with the CI for the empirical chance level. All multivariate predictions reported in **Table 2** were determined on a 5-fold cross-validation to evaluate the model's generalization to new data to approximate an online analysis.

This is to our knowledge the first evidence that brain processes may differ in the interactions between human driven and autonomous cars. Together with the behavioral results, this suggests that human driver may assess the interactions with AV differently from interactions with HV.

Effect Sizes Discriminating Turning in Front of AV vs. HV From fNIRS Brain Activation

Figure 5 shows the Cohen's d brain maps for the group-level analysis. We visualized the averaged brain map on the MNI 152 brain in Neurosynth[1] and used MRIcron[2] to determine MNI coordinates and the corresponding Brodmann areas (BA) for the brain areas with increased activation differences during the left-lane turning decision-making phases for AVs as compared to HVs.

Table 3 lists the brain areas, the MNI-coordinates of the difference maxima and the Cohen's d values as indicators of the effect sizes from the group-level analyses ($n = 12$).

The results showed the largest effect sizes of brain activation in the prefrontal cortex (PFC), reflecting activation changes in the left and right dorsolateral areas (dlPFC; putative BA 46) and the left ventrolateral prefrontal (vlPFC; putative BA 45) areas (Cohen's d ∼ 0.9–1.2). Additionally, the ventromedial prefrontal areas (vmPFC; putative BA 10) also indicate increased activation differences while turning in front of AV as compared to HV (Cohen's d ∼ 0.9). These prefrontal areas have been previously implicated in the valuation of actions during decision-making (Sanfey, 2007; Lee, 2008; Rangel et al., 2008; Hollmann et al., 2011; Ruff and Fehr, 2014). Furthermore, the superior frontal gyrus (SFG) and parts of the motor cortices (putative BA 6) also show increased activation differences between the turning phases of AV and HV. Moreover, some informative channels can be seen in the left superior parietal areas (putative BA 7). Overall, our

TABLE 2 | Five-fold cross-validated predictions of AV and HV from HbR fNIRS measurements using multivariate logistic ridge regression analysis for all participants.

Participant	1	2	4	5	6	7	8	9	10	11	12	13	Mean
Accuracy in % (SD)	69.9 (0.04)	56.8 (0.04)	67.9 (0.03)	68.5 (0.03)	59.5 (0.05)	68.1 (0.04)	65.4 (0.03)	68.7 (0.04)	66.0 (0.04)	66.3 (0.04)	75.9 (0.04)	73.7 (0.05)	67.2 (0.05)
F1-score (SD)	0.70 (0.04)	0.57 (0.02)	0.67 (0.03)	0.69 (0.03)	0.61 (0.05)	0.69 (0.04)	0.65 (0.03)	0.69 (0.04)	0.66 (0.04)	0.64 (0.04)	0.77 (0.04)	0.74 (0.05)	0.67 (0.05)
5–95% accuracy CI for empirical chance level (SD)	45.8–54.1	46.9–53.1	46.2–53.8	46.2–53.5	46.3–54.1	46.0–54.2	46.6–53.4	45.6–54.3	46.0–53.6	46.2–53.9	45.3–54.2	45.6–54.5	46.0–53.9 (0.004)

[1]http://neurosynth.org
[2]https://www.nitrc.org/projects/mricron

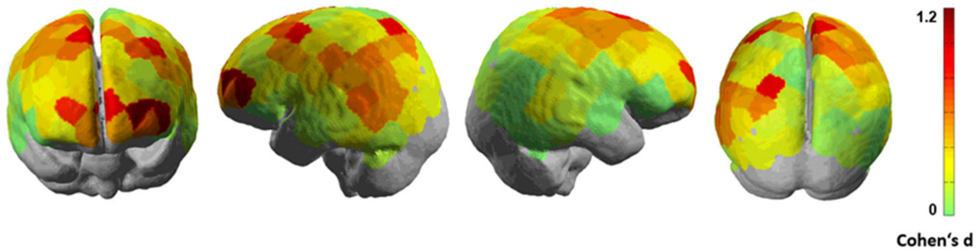


FIGURE 5 | Cohen's d brain maps representing effect sizes computed from channel-wise weighted averaged t-statistics (t_{avg}) for the group-level analysis. Moderate to high Cohen's d values (0.8–1.2) show medium to large effect sizes indicating increased activation differences for the decision-making phase during turning in front of AV as compared to HV.

TABLE 3 | Brain areas showing increased activation differences in the decision-making phase during turning in front of AVs compared to HVs.

Brain areas	Putative Brodmann area (BA)	X	Y	Z	Cohen's d
Left dorsolateral prefrontal	46	−26	62	26	1.20
Right dorsolateral prefrontal	46	24	60	32	0.81
Ventromedial prefrontal	10	4	58	28	0.93
Left ventrolateral prefrontal	45	−48	42	28	1.10
Left superior frontal gyrus	6	−34	10	60	0.87
Right superior frontal gyrus	6	16	4	76	0.93
Left superior parietal	7	−28	−62	68	0.86

The approximate MNI coordinates of activation differences along with the putative Brodmann areas and their Cohen's d values are shown.

results demonstrate a consistent difference in activation at the brain-level and these activation differences occur in brain areas that have been previously related to decision-making.

DISCUSSION

The main promise of autonomous driving is that AVs will reduce traffic accidents caused by human errors and hence be safer than HVs. The aim of this study was to investigate if there is a difference between the valuation of actions when an interaction involves technical systems such as AVs as compared to similar interactions with other HVs. Moreover, we wanted to investigate if these potential differences in human-human and human-autonomous vehicle interactions can be characterized from behavior and neurophysiological whole-head fNIRS brain activation measurements. We believe that this research goal is extremely relevant in the present situation since some studies have shown that humans could actively exploit the predictable and safe behavior of AVs. With the knowledge that AVs are programmed to be risk-averse, humans tend to act with impunity while interacting with AVs (Millard-Ball, 2018; Trende et al., 2019; Liu et al., 2020). Our results provide evidence that humans show a difference in the valuation of actions in the decisions they make in such situations depending on whether they interact with an AV or a HV and this is expressed in fNIRS brain activation and partly in the behavioral tendencies.

We investigated differences in human-human and human-autonomous interactions using a full-scale fixed base driving. In our cover story, we mentioned that the AVs were defensively programmed in an interaction and drove conservatively to avoid collisions and had faster braking reaction times than HVs as this is the expected programming of AVs (Zhan et al., 2016; Li and Sun, 2018). However, both, AVs and HVs were simulated according to the same driving behavior.

Results of our gap acceptance model showed that the confidence intervals of the threshold parameter (m) of the gap acceptance models overlapped, indicating that there is no significant difference between the safety margin used during the turning maneuvers with respect to AVs or HVs. Furthermore, we observed differences in the model widths, which describes the 5–95% point of the AV and HV models. The models' width parameters indicated that the AV distribution is steeper than the HV distribution. The steeper slope for AV could be interpreted as participants tended to be more certain in their decision-making process while turning in front of AV as reflected by the smaller variance in gap size acceptance as compared to HV. Similarly, the shallower slope for HV could indicate that participants have larger uncertainty in their decision-making process while interacting with HV during the lane-turning maneuver. We assume that the participants may have felt more certain while interacting with AVs due to their perceived predictability and potentially tried to exploit the defensive strategy of AVs. The participants may have overestimated the AVs' alleged defensive behavior despite the fact that AVs showed the same driving behavior as HVs in order to gain a temporal advantage in the experiment. This assumption is further supported by the results of a trust questionnaire. The mean score for the item "I trust AVs more than HVs" was 3.9 (on a scale of 1–6) supporting the claim that the participants may have overestimated the AVs' alleged driving behavior and underestimated the technological limitations of AVs. This automation complacency regarding the AVs' safe functioning in the simulation could potentially lead to dangerous situations (Parasuraman and Riley, 1997).

The neurophysiological results indicate that our approach of using whole-head fNIRS in combination with a cross-validated multivariate logistic ridge regression is suitable to predict the type of involved traffic agent (AV or HV) while making a decision to turn. This approach allowed us to exploit the spatial specificity of whole-head fNIRS, in order to predict the traffic agent involved at the crossing with an average accuracy of approximately 67% (SD = 3%) across all participants and up to a maximum of almost 76% on a single-subject level. It is important to note that these cross-validated predictions are obtained from just 4 s of fNIRS data in the decision-making phase demonstrating that our approach of combining multivariate logistic ridge regression and cross-validation and exploiting the spatial specificity of whole-head fNIRS has the potential to predict the interaction partner in time-critical situations. While the predictions might not be very high, we have previously shown that even imperfect predictions regarding the driver's intent can be useful to develop driver models which can lead to increased safety during interactions between AVs and HVs in mixed traffic environments (Damm et al., 2019).

To characterize the neural correlates for the decision-making phase in human-human and human-autonomous interactions, we computed the channel-wise Cohen's d measures as effect sizes for the fNIRS brain activation over the AV and HV turning conditions in a group-level analysis. Our initial hypothesis was that human-autonomous vehicle interaction would result in increased modulations in the prefrontal areas such as the dorsal and ventral frontal areas, frontal midline brain areas such as the ventromedial prefrontal areas and the anterior cingulate cortex, since these brain areas are thought to represent the values of actions taken (Sanfey, 2007; Rangel et al., 2008; Ruff and Fehr, 2014). Due to the limited spatial depth of fNIRS, we could not observe the activity in the anterior cingulate cortex. However, the group-level analysis revealed increased fNIRS activation in the prefrontal areas such as the dorsolateral (putative BA 46), left ventrolateral (putative BA 45), ventromedial prefrontal areas (putative BA 10), the superior frontal gyrus and parts of the motor cortices (putative BA 6) when participants turned in front of the AV as compared to HV. The activation in these brain areas could potentially reflect the differences in valuation of actions when turning in front of an AV as compared to HV. The prefrontal cortex is an important brain area that subserves higher order executive functions necessary for the cognitive control of behavior and decision-making. The dorsolateral prefrontal areas (putative BA 46) show increased activation during risky decision-making where costs and benefits are weighed (Duncan et al., 1996). BA 45 has been associated with reasoning and goal-intensive processing (Goel et al., 1998; Fincham et al., 2002). The ventromedial prefrontal cortex (putative BA 10) has been shown to be a part of the reward-processing mechanism elicited by emotional processes, which plays a vital role in determining value-based decision-making (Sanfey, 2007). Moreover, some studies have shown the role of the dorsolateral and ventromedial prefrontal areas to be involved in uncertainty during the decision-making processes (Schienle et al., 2010; Stern et al., 2010; Wever et al., 2015; Tomov et al., 2020). This can be linked to our interpretation of our behavioral results which show a difference in the certainty of the driver during the planning and execution of the turning maneuver in the decision-making phase while interacting with AVs or HVs.

Previous studies have shown the role of the superior frontal gyrus in processing emotions and self-reflections in decision-making (Deppe et al., 2005; Goldberg et al., 2006). Additionally, the involvement of the BA 6 in motor functioning such as planning and execution of motor activities is well-known (Catalan et al., 1998; Hanakawa et al., 2008) suggesting differences in the underlying brain processes during interactions with AVs and HVs.

In our study, some participants (7 out of 12) mentioned that they deliberately took the gap in front of the AVs because they assumed it would brake due to the alleged defensive behavior. This is a dangerous assumption since all the vehicles in the simulation including HVs and AVs were simulated according to the same driving behavior. The participants overestimated the behavior and driving performance of the vehicles solely based on the cover story about the defensive programming of the vehicles and ignored the visual evidence based on the similar driving behavior of AVs and HVs. This is a classic example of "misuse of automation" as defined by Parasuraman and Riley (1997) as an overreliance of automation. This automation complacency may lead to dangerous traffic situations or even accidents in case of excessive overestimation of the reaction time of the AV or sensor failure (Parasuraman and Manzey, 2010). Most of the participants in this study believed that AVs are safer. The findings of this study may be important in mixed traffic environments where both HVs and AVs are participating in the traffic. The software controlling AVs should be able to account for the fact that humans may behave riskier during interactions. Furthermore, it would be interesting to investigate how human drivers would behave if the AVs were able to retaliate uncooperative or risky driving behavior by providing clearly visible cues. Future studies could investigate if the behavior of the human driver changes and if this is reflected in a change of the action valuation signals in the brain activation becoming more similar to the activations observed in interactions with HVs.

The current study has a few limitations. The experimental design of the gap sizes did not feature sufficient samples with gap sizes in the range of 3–6 s. This leads to fewer events within this range. Furthermore, it should be mentioned that the experiment was conducted with a rather homogenous participant pool. The participants were mainly between 21 and 29 years and from an academic background. This group is generally associated to have high trust in technology (Kennedy et al., 2008) which may have an impact on the results from the subjective questionnaire regarding high trust in AV. We suspect that participants with low trust in technology in general and less trust in the safe functioning of AVs in particular will behave differently in such an experiment leading to a shallower slope in gap acceptance function for AVs. In this study, the only relevant factors for the gap acceptance model were the gap size and traffic agent involved, i.e., HVs or AVs. Several studies argue that the gap acceptance also depends on personal characteristics such as age, gender, or intersection characteristics (Darzentas et al., 1980; Bottom and Ashworth, 2007; Yan et al., 2007) which have not been considered in this study.

The brain areas characterized in this study have been shown to be involved in determining the valuation of actions during social interactions in lab-based settings. These neural correlates could be used to develop control systems for interactions with AVs at intersections based on the behavioral tendencies of the driver. Moreover, neurophysiological measures could be used as an indicator to predict the intent of the driver in such human-autonomous interactions. Furthermore, integrating such neurophysiological sensors in control systems could potentially optimize the performance of the AVs under safety constraints in mixed traffic environments in the future.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Carl von Ossietzky University, Oldenburg. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AL, KB, AP, MF, and JR planned the research. AT and LW developed the experimental paradigm with support from BB and SK. AU, AT, and CP carried out the data collection and the data analysis. AU and AT prepared the manuscript. All authors provided feedback. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under grant numbers RI 1511/3-1 to JR, LU 1880/3-1 to AL, and BE 4532/15-1 to KB (Learning from Humans—Building for Humans), PR 1266/3-1 to AP (Design Paradigms for Societal-Scale Cyber-Physical Systems), and FR 2715/4-1 (Integrated Socio-technical Models for Conflict Resolution and Causal Reasoning) to MF.

REFERENCES

Biebl, B., and Bengler, K. (2021). I spy with my mental eye – analyzing compensatory scanning in drivers with homonymous visual field loss. *Lect. Notes Networks Syst.* 221, 552–559. doi: 10.1007/978-3-030-74608-7_67

Bottom, C. G., and Ashworth, R. (2007). Factors affecting the variability of driver gap-acceptance behaviour. *Psychol. Aging* 21, 721–734. doi: 10.1080/00140137808931774

Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., et al. (2014). Motion artifacts in functional near-infrared spectroscopy: a comparison of motion correction techniques applied to real cognitive data. *NeuroImage* 85, 181–191. doi: 10.1016/j.neuroimage.2013.04.082

Catalan, M. J., Honda, M., Weeks, R. A., Cohen, L. G., and Hallett, M. (1998). The functional neuroanatomy of simple and complex sequential finger movements: a PET study. *Brain* 121, 253–264. doi: 10.1093/brain/121.2.253

Choi, E.-H. (2010). *Crash Factors in Intersection-Related Crashes: An On-Scene Perspective*. Createspace Independent Pub. Available online at: https://books.google.de/books?id=Pxc2nQAACAAJ

Cunningham, M., and Regan, M. A. (2015). *Autonomous Vehicles: Human Factors Issues and Future Research.* Australian College of Road Safety.

Damm, W., Fränzle, M., Lüdtke, A., Rieger, J. W., Trende, A., and Unni, A. (2019). Integrating neurophysiological sensors and driver models for safe and performant automated vehicle control in mixed traffic. *IEEE Intelligent Vehicles Symp. Proc.* 2019, 82–89. doi: 10.1109/IVS.2019.8814188

Darzentas, J., McDowell, M., and Cooper, D. (1980). Minimum acceptable gaps and conflict involvement in a simple crossing manuevre. *Traffic Eng Control* 21, 58–61.

Deppe, M., Schwindt, W., Kugel, H., Plassmann, H., and Kenning, P. (2005). Nonlinear responses within the medial prefrontal cortex reveal when specific implicit information influences economic decision making. *J. Neuroimaging* 15, 171–182. doi: 10.1177/1051228405275074

Duncan, J., Emslie, H., Williams, P., Johnson, R., and Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cogn. Psychol.* 30, 257–303. doi: 10.1006/cogp.1996.0008

Fincham, J. M., Carter, C. S., Van Veen, V., Stenger, V. A., and Anderson, J. R. (2002). Neural mechanisms of planning: A computational analysis using event-related fMRI. *Proc. Natl. Acad. Sci. U.S.A.* 99, 3346–3351. doi: 10.1073/pnas.052703399

Fitzpatrick, K. (1991). *Gaps Accepted at Stop-Controlled Intersections*. Transportation Research Record no. 1303. Transportation Research Board. Available online at: https://trid.trb.org/view/359881

Foy, H. J., Runham, P., and Chapman, P. (2016). Prefrontal cortex activation and young driver behaviour: A fNIRS study. *PLoS ONE* 11, e0156512. doi: 10.1371/journal.pone.0156512

Gerstenberger, M. (2015). *Unfallgeschehen an Knotenpunkten. In Dissertation*. Available online at: https://katalog.ub.tu-braunschweig.de/vufind/Record/834200899 (accessed December 14, 2021).

Gläscher, J., Hampton, A. N., and O'Doherty, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb. Cortex* 19, 483–495.

Goel, V., Gold, B., Kapur, S., and Houle, S. (1998). Neuroanatomical correlates of human reasoning. *J. Cogn. Neurosci.* 10, 293–302. doi: 10.1162/089892998562744

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., et al. (2006). The international personality item pool and the future of public-domain personality measures. *J. Res. Personal.* 40, 84. doi: 10.1016/j.jrp.2005.08.007

Hanakawa, T., Dimyan, M. A., and Hallett, M. (2008). Motor planning, imagery, and execution in the distributed motor network: a time-course study with functional MRI. *Cereb. Cortex* 18, 2775–2788. doi: 10.1093/CERCOR/BHN036

Hastie, T., Tibshirani, R., and Friedman, J. H. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.), New York, NY: Springer.

Hollmann, M., Rieger, J. W., Baecke, S., Lützkendorf, R., Müller, C., Adolf, D., et al. (2011). Predicting decisions in human social interactions using real-time fMRI and pattern classification. *PLoS ONE* 6, e25304. doi: 10.1371/journal.pone.0025304

Huppert, T. J., Franceschini, M. A., and Boas, D. A. (2009). "Noninvasive imaging of cerebral activation with diffuse optical tomography", in *In Vivo Optical*

Imaging of Brain Function (2nd ed.), ed. R. D. Frostig (Boca Raton, FL: CRC Press/Taylor & Francis). Available online at: http://europepmc.org/books/NBK20225

Huppert, T. J., Hoge, R. D., Diamond, S. G., Franceschini, M. A., and Boas, D. A. (2006). A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *NeuroImage* 29, 368–382. doi: 10.1016/j.neuroimage.2005.08.065

Jolliffe, I. T. (1972). Discarding variables in a principal component analysis. I: Artificial data. *Appl. Statistics* 1972, 160–173.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23, 187–200.

Kennedy, G. E., Judd, T. S., Churchward, A., Gray, K., and Krause, K. L. (2008). First year students' experiences with technology: Are they really digital natives? *Austr. J. Educ. Technol.* 24, 108–122. doi: 10.14742/ajet.1233

Krueger, H.-P., Grein, M., Kaussner, A., and Mark, C. (2005). *SILAB-A Task Oriented Driving Simulation*.

Kurpiers, C., Biebl, B., Hernandez, J. M., and Raisch, F. (2020). Mode awareness and automated driving—what is it and how can it be measured? *Information* 11, 277. doi: 10.3390/INFO11050277

Lee, D. (2008). Game theory and neural basis of social decision making. *Nat. Neurosci.* 11, 404–409. doi: 10.1038/nn2065

Li, X., and Sun, J. Q. (2018). Defensive driving strategy and control for autonomous ground vehicle in mixed traffic. *Stud. Comput. Intelligence* 731, 3–44. doi: 10.1007/978-3-319-64063-1_1

Liu, P., Du, Y., Wang, L., and Da Young, J. (2020). Ready to bully automated vehicles on public roads? *Accident Anal. Prevent.* 137, 105457. doi: 10.1016/J.AAP.2020.105457

Lord-Attivor, R., and Jha, M. K. (2012). "Modeling gap acceptance and driver behavior at stop controlled (priority) intersections in developing countries," *Proceedings of The 6th WSEAS International Conference on Computer Engineering and Applications*. pp. 29–38.

MacIntosh, B. J., Klassen, L. M., and Menon, R. S. (2003). Transient hemodynamics during a breath hold challenge in a two part functional imaging study with simultaneous near-infrared spectroscopy in adult humans. *NeuroImage* 20, 1246–1252. doi: 10.1016/S1053-8119(03)00417-8

Millard-Ball, A. (2018). Pedestrians, autonomous vehicles, and cities. *J. Plan. Educ. Res.* 38, 6–12. doi: 10.1177/0739456X16675674

Obrig, H., Neufang, M., Wenzel, R., Kohl, M., Steinbrink, J., Einhäupl, K., et al. (2000). Spontaneous low frequency oscillations of cerebral hemodynamics and metabolism in human adults. *NeuroImage* 12, 623–639. doi: 10.1006/nimg.2000.0657

Parasuraman, R., and Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Hum. Fact.* 52, 381–410. doi: 10.1177/0018720810376055

Parasuraman, R., and Riley, V. (1997). Humans and automation: use, misuse, disuse, abuse. *Hum. Fact.* 39, 230–253. doi: 10.1518/001872097778543886

Qian, J., Hastie, T., Friedman, J., Tibshirani, R., and Simon, N. (2013). *Glmnet for Matlab 2013*. Available online at: http://www.stanford.edu/~hastie/glmnet_matlab (accessed December 14, 2021).

Ragland, D. R., Arroyo, S., and Shladover, S. E. (2006). *UC Berkeley Research Reports Title Gap Acceptance for Vehicles Turning Left Across On-Coming Traffic: Implications for Intersection Decision Support design Permalink*. Available online at: https://escholarship.org/uc/item/8455h5gq (accessed December 14, 2021).

Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556. doi: 10.1038/nrn2357

Reichert, C., Fendrich, R., Bernarding, J., Tempelmann, C., Hinrichs, H., and Rieger, J. W. (2014). Online tracking of the contents of conscious perception using real-time fMRI. *Front. Neurosci.* 8, 116. doi: 10.3389/fnins.2014.00116

Ruff, C. C., and Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nat. Rev. Neurosci.* 15, 549–562. doi: 10.1038/nrn3776

Saffarian, M., De Winter, J. C. F., and Happee, R. (2012). Automated driving: human-factors issues and design solutions. *Sage J.* 2012, 2296–2300. doi: 10.1177/1071181312561483

Sanfey, A. G. (2007). Social decision-making: Insights from game theory and neuroscience. *Science* 318, 598–602. doi: 10.1126/science.1142996

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science* 300, 1755–1758. doi: 10.1126/science.1082976

Sassaroli, A., and Fantini, S. (2004). Comment on the modified Beer-Lambert law for scattering media. *Phys. Med. Biol.* 49, N255–N257.

Schienle, A., Köchel, A., Ebner, F., Reishofer, G., and Schäfer, A. (2010). Neural correlates of intolerance of uncertainty. *Neurosci. Lett.* 479, 272–276. doi: 10.1016/J.NEULET.2010.05.078

Schmitz, C. H., Klemer, D. P., Hardin, R., Katz, M. S., Pei, Y., Graber, H. L., et al. (2005). Design and implementation of dynamic near-infrared optical tomographic imaging instrumentation for simultaneous dual-breast measurements. *Appl. Optics* 44, 2140. doi: 10.1364/AO.44.002140

Schneider, P., Piper, S., Schmitz, C., Schreiter, N., Volkwein, N., Lüdemann, L., et al. (2011). Fast 3D Near-Infrared Breast Imaging Using Indocyanine Green for Detection and Characterization of Breast Lesions. *RöFo* 183, 956–963. doi: 10.1055/s-0031-1281726

Schroeter, M. L., Kupka, T., Mildner, T., Uludag, K., and von Cramon, D. Y. (2006). Investigating the post-stimulus undershoot of the BOLD signal—a simultaneous fMRI and fNIRS study. *NeuroImage* 30, 349–358. doi: 10.1016/j.neuroimage.2005.09.048

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron* 36, 241–263. doi: 10.1016/S0896-6273(02)00967-4

Schütt, H., Harmeling, S., Macke, J., and Wichmann, F. (2015). Psignifit 4: pain-free bayesian inference for psychometric functions. *J. Vision* 15, 474–474. doi: 10.1167/15.12.474

Singh, S. (2018). *Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey*. Traffic Safety Facts - Crash Stats. Available online at: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812506 (accessed February 10, 2022).

Stern, E. R., Gonzalez, R., Welsh, R. C., and Taylor, S. F. (2010). Updating beliefs for a decision: neural correlates of uncertainty and underconfidence. *J. Neurosci.* 30, 8032–8041. doi: 10.1523/JNEUROSCI.4729-09.2010

Suzuki, S. (2017). Optimized statistical parametric mapping procedure for NIRS data contaminated by motion artifacts : Neurometric analysis of body schema extension. *Brain Inform.* 4, 171–182. doi: 10.1007/s40708-017-0070-x

Taatgen, N., Lebiere, C., and Anderson, J. (2005). *Modeling Paradigms in ACT-R. Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511610721.003

Tay, R., and Rifaat, S. M. (2007). Factors contributing to the severity of intersection crashes. *J. Adv. Transport.* 41, 245–265. doi: 10.1002/atr.5670410303

Tomov, M. S., Truong, V. Q., Hundia, R. A., and Gershman, S. J. (2020). Dissociable neural correlates of uncertainty underlie different exploration strategies. *Nat. Commun.* 11, 15766. doi: 10.1038/S41467-020-15766-Z

Trende, A., Unni, A., Weber, L., Rieger, J. W., and Luedtke, A. (2019). An investigation into human-autonomous vs. Human-human vehicle interaction in time-critical situations. *ACM Int. Conf. Proc. Series* 2019, 303–304. doi: 10.1145/3316782.3321544

Unni, A., Ihme, K., Jipp, M., and Rieger, J. W. (2017). Assessing the driver's current level of working memory load with high density functional near-infrared spectroscopy: a realistic driving simulator study. *Front. Hum. Neurosci.* 11, 167. doi: 10.3389/FNHUM.2017.00167

Villringer, A., Planck, J., Hock, C., Schleinkofer, L., and Dirnagl, U. (1993). Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults. *Neurosci. Lett.* 154, 101–104. doi: 10.1016/0304-3940(93)90181-J

Vollrath, M., Briest, S., Schießl, C., Drews, J., and Becker, U. (2006). *Ableitung von Anforderungen an Fahrerassistenzsysteme aus Sicht der Verkehrssicherheit*. Available online at: https://www.bast.de/DE/Publikationen/Berichte/unterreihe-f/2007-2000/f60.html (accessed December 14, 2021).

Wever, M., Smeets, P., and Sternheim, L. (2015). Neural correlates of intolerance of uncertainty in clinical disorders. *J. Neuropsychiatry Clin. Neurosci.* 27, 345–353. doi: 10.1176/appi.neuropsych.14120387

Xu, Y., Graber, H. L., and Barbour, R. L. (2014). nirsLAB: A computing environment for fNIRS neuroimaging data analysis. *Biomed. Optics* 2014, BM3A.1. doi: 10.1364/BIOMED.2014.BM3A.1

Yan, X., Radwan, E., and Guo, D. (2007). Effects of major-road vehicle speed and driver age and gender on left-turn gap acceptance. *Acc. Anal. Preven.* 39, 843–852. doi: 10.1016/j.aap.2006.12.006

Zhan, W., Liu, C., Chan, C. Y., and Tomizuka, M. (2016). A non-conservatively defensive strategy for urban autonomous driving. *IEEE Conf. Intel. Transport. Syst. Proc.* 2016, 459–464. doi: 10.1109/ITSC.2016.7795595

Zhang, Q., Strangman, G. E., and Ganis, G. (2009). Adaptive filtering to reduce global interference in non-invasive NIRS measures of brain activation: How well and when does it work? *NeuroImage* 45, 788–794. doi: 10.1016/j.neuroimage.2008.12.048

Zhang, Y., Brooks, D. H., Franceschini, M. A., and Boas, D. A. (2005). Eigenvector-based spatial filtering for reduction of physiological interference in diffuse optical imaging. *J. Biomed. Optics* 10, 011014. doi: 10.1117/1.1852552

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

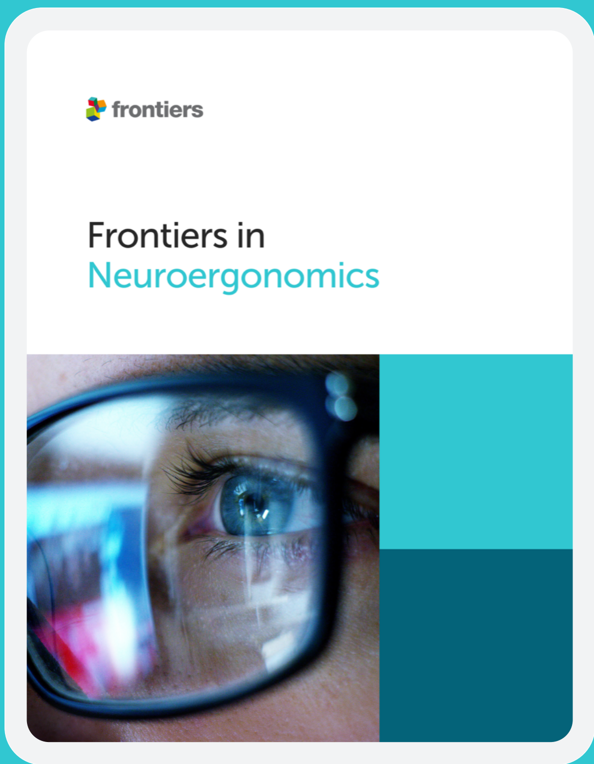
Frontiers in Neuroergonomics

Using neuroscience to understand the brain in everyday life

Part of the most cited neuroscience series, this journal advances neuroscience and neuroengineering applications to human factors, ergonomics and engineering, sharing new solutions for understanding the human brain at work.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact