THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Effects of cognitive tasks on car drivers' behaviors and physiological responses

EMMA NILSSON

Department of Mechanics and Maritime Sciences

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2022

Effects of cognitive tasks on car drivers' behaviors and physiological responses
EMMA NILSSON
ISBN 978-91-7905-710-7

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5176
ISSN 0346-718X

Department of Mechanics and Maritime Sciences
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

The cover image depicts an emoji driving a car while engaged in a cognitive task. The neutral emoji is downloaded from https://emojiisland.com and has been modified by the author.

Chalmers digitaltryck
Gothenburg, Sweden 2022

Effects of cognitive tasks on car drivers' behaviors and physiological responses

EMMA NILSSON

Department of Mechanics and Maritime Sciences
Chalmers University of Technology

ABSTRACT

The effects of drivers' engagement in cognitive tasks (i.e., non-visual, cognitively loading activities unrelated to the task of driving) are debated and unclear. Numerous experiments show impaired driver behaviors, yet naturalistic studies typically do not find an increased crash risk. In the future, autonomous driving (AD) is expected to improve traffic safety while allowing safe engagement in cognitive (and other) tasks. Having the opportunity to perform non-driving related tasks while traveling may then motivate drivers to use AD, provided they can actually engage in the tasks. Unfortunately, research on drivers' engagement in cognitive tasks suffers severe methodological limitations since reliable and unintrusive measures of cognitive load are lacking.

The aim of this thesis is therefore to advance the understanding of task-induced cognitive load in the context of traffic safety. This aim is split into two objectives: A) to better understand how drivers' involvement in cognitive tasks can affect safety-relevant driver behaviors and decisions and B) to provide methodological guidance about assessing cognitive load in drivers using physiological measures.

To accomplish Objective A, effects of cognitive tasks on driver behaviors were studied during routine driving and in a safety-critical event in a driving simulator. Also, drivers' ability to engage in a non-driving related task while using AD in real traffic was explored. In line with the cognitive control hypothesis (Engström et al., 2017), it was found that cognitive tasks negatively affected driver behaviors in situations where cognitive control was needed, for example in intersections—but not in a lead vehicle braking scenario where responses were triggered automatically by visual looming. It was also found that although the number of off-path glances decreased during cognitive load, the timing of the remaining glances was unaffected. Clearly, cognitive load has different effects on different mechanisms. When using AD, drivers were indeed capable of engaging in a non-driving related task—suggesting that AD will be able to fulfill drivers' desire to perform such tasks while traveling, which may motivate AD usage and thus improve traffic safety (given that AD is truly safer than manual driving). Finally, a simulator study addressing Objective B showed that the measurability of cognitive load was greatly improved by recognizing that multiple coexisting mental responses give rise to different physiological responses. This approach can provide less context-dependent measurements and allows for a better, more detailed understanding of the effects of cognitive tasks.

These findings can help improve traffic safety—both by being used in system development, and as part of the systems themselves.

Key words: cognitive load, distraction, physiological measures, driver behavior, traffic safety, inattention, psychophysiology, autonomous drive.

ACKNOWLEDGMENTS

Much of this work was carried out as parts of the projects Vehicle Driver Monitoring (VDM) and HARMONISE. Both projects were supported by grants from the Swedish governmental agency for innovation (VINNOVA). The remainder of the work was funded by Volvo Car Corporation.

There are many people who have supported me in various ways in my PhD work and I am very thankful to all of you. My supervisors (current and former): Mikael Ljung Aust, Jonas Bärgman, Trent Victor, and Rob Thomson—thank you for all the great guidance and support over these years. All my colleagues in VDM and HARMONISE—you gave me new perspectives, taught me a lot, and I really enjoyed working with you all. A special thanks to Bo Svanberg for your never-ending engagement and great friendship, and to Per Lindén for all your help and patience. I would also like to thank Henrik Wiberg—without you, this PhD work would most likely not have happened.

I would also like to thank my family; Fredrik—for your love, support, and patience, Majken, Vera, and Sigrid—for being so wonderful, and last but certainly not least, my mother Lotta and my mother-in-law Sonja—thank you so much for your tremendous help with the kids, I could not have done this without you!

APPENDED PUBLICATIONS

Paper I Nilsson, E. J., Ljung Aust, M., Engström, J., Svanberg, B., & Lindén, P. (2018). Effects of cognitive load on response time in an unexpected lead vehicle braking scenario and the detection response task (DRT). *Transportation Research Part F, 59*, 463-474. https://doi.org/10.1016/j.trf.2018.09.026

Contribution: Nilsson contributed to the study design, performed the formal analysis, and wrote most of the original draft of the paper.

Paper II Nilsson, E. J., Victor, T., Ljung Aust, M., Svanberg, B., Lindén, P., & Gustavsson, P. (2020). On-to-off-path gaze shift cancellations lead to gaze concentration in cognitively loaded car drivers: A simulator study exploring gaze patterns in relation to a cognitive task and the traffic environment. *Transportation Research Part F, 75*, 1-15. https://doi.org/10.1016/j.trf.2020.09.013

Contribution: Nilsson did most of the study design, performed the formal analysis and visualizations, and wrote the original draft of the paper.

Paper III Klingegård, M., Andersson, J., Habibovic, A., Nilsson, E., & Rydström, A. (2020). Drivers' ability to engage in a non-driving related task while in automated driving mode in real traffic. *IEEE Access, 8*, 221654-221668. https://doi.org/10.1109/ACCESS.2020.3043428

Contribution: Nilsson participated in the design of the study, contributed to the data curation of glance data, and did the formal analysis and visualizations that involved glance data. She wrote a few sections of the original draft of the paper and participated in the editing and review process of the whole paper.

Paper IV Nilsson, E. J., Bärgman, J., Ljung Aust, M., Matthews, G., & Svanberg, B. (2022). Let complexity bring clarity: A multidimensional assessment of cognitive load using physiological measures. *Frontiers in Neuroergonomics, 3*, Article 787295. https://doi.org/10.3389/fnrgo.2022.787295

Contribution: Nilsson did most of the study design and processing of the physiological data. She performed the formal analysis and visualizations, and she wrote the original draft of the paper.

CONTENTS

1 Introduction

As vehicles' technical development progresses, great improvements in traffic safety are possible. This is indeed needed, since road injuries are one of the ten leading causes of death worldwide and the leading cause of death for people aged 5–29 years (World Health Organization [WHO], 2018). Consequently, one of the United Nations' sustainable development goals is to halve the number of deaths and injuries from road traffic accidents (United Nations [UN], 2015). Today, driver errors are considered the most common reason for traffic accidents (Deme, 2019; Singh, 2015). Therefore, knowledge about drivers' abilities, behaviors, priorities, and experiences is needed to develop vehicles and vehicle applications which will improve traffic safety by reducing these errors.

In routine driving (i.e., typical "everyday driving" without critical events), drivers need to attend to relevant information in the traffic environment and adapt their behavior accordingly. For example, if a bus has stopped at a bus stop, an aware driver would recognize the risk that a pedestrian might be hidden by the bus and may be about to appear on the road. The driver should thus direct his/her gaze towards that location, slow down if needed, and be ready to brake immediately if a pedestrian appears. Also, if drivers fail to adapt to the traffic situation—or confront an unpredictable or unavoidable critical event—they need to respond quickly and appropriately to avoid a crash (or, if a crash is inevitable, to limit its damage).

Drivers' engagement in non-driving-related activities has been frequently studied for its potential negative effects on these adaptation and response abilities and thus on traffic safety. Various studies have repeatedly and consistently shown that visual distraction (i.e., not looking at driving-relevant targets; Ahlström et al., 2012) has detrimental effects on driver behaviors and traffic safety (Angell et al., 2006; Caird et al., 2014). Clearly, diverting the gaze from the driving task decreases the chance of noticing relevant information in the traffic environment and responding promptly in case of an impending threat (Drews et al., 2009; Liang & Lee, 2010; Victor et al., 2015), and it increases the risk of crashes and near-crashes (Fitch et al., 2013; Klauer et al., 2006). This finding has led many countries to introduce laws to reduce visual distraction by, for example, banning handheld mobile phone use while driving (WHO, 2018). Technical solutions have also been developed, such as voice-based auditory interfaces for in-vehicle and portable devices, which are intended to minimize the devices' negative effects on traffic safety by enabling drivers to handle the devices while keeping their eyes on the road. However, strong concerns have been raised regarding negative effects also of these non-visual tasks. But unlike the research results for visual distraction, the results of the effects of these tasks are ambiguous. I will refer hereafter to these non-visual, cognitively loading, non-driving-related tasks as cognitive tasks, to differentiate them from tasks such as texting, which include a visual component (thus avoiding the specific but awkward acronym NVCLNDRT).

On one hand, numerous studies have reported negative effects on driver performance of engagement in cognitive tasks such as cell phone conversations and working memory loading tasks (for reviews, see Collet et al., 2010; Caird et al., 2018; von Janczewski et al., 2021). The effects include deteriorated visual scanning (Sawyer et al., 2016), hazard detection (G. Wood et al., 2016), and situational adaptation (Baumann et al., 2008; Muttart et al., 2007),

and increased response times to various stimuli (e.g., Strayer et al., 2013). These effects, found in controlled driving experiments, are assumed to generalize to real-life driving, leading to predictions of an increase in safety-critical events due to drivers' missing relevant information in the traffic environment (Recarte & Nunes, 2003), and an increased risk of crashes due to slower responses in critical situations (Collet et al., 2010; Strayer et al., 2015; Merat & Jamson, 2008).

On the other hand, contrary to these inferences, naturalistic studies (i.e., studies with instrumented vehicles in everyday driving in real traffic) typically haven't found an increased risk of crashes or near-crashes during cognitive tasks such as cell phone conversations or the use of CB (Citizens Band) radios (Fitch et al., 2013; Klauer et al., 2006). Instead, several studies have even found a decreased risk of crashes and near-crashes during such tasks (Hickman et al., 2010; Olson et al., 2009; Victor et al., 2015). These results make some researchers question the conclusions drawn from controlled experiments (Engström et al., 2017; Fisher, 2015; Hancock & Sawyer, 2015; Shinar, 2015). Numerous explanations for the seemingly discrepant findings have been suggested (see Wijayaratna et al., 2019, for a review). For example, drivers' engagement in cognitive tasks in real-life driving may help them stay alert (Olson et al., 2009) and keep them from engaging in more safety-detrimental activities such as texting (Victor et al., 2015). It has also been suggested that drivers adapt their behavior to counteract the negative effects of task engagement (Törnros & Bolling, 2005), and that the differences in the effects of cognitive tasks between controlled experiments and real-life driving are due to differences in the driving tasks (Engström et al., 2017; Shinar, 2015).

To resolve the many discrepancies between studies and understand how cognitive tasks affect traffic safety, knowledge is needed about how, when, and where drivers engage in cognitive tasks (see e.g., Tivesten & Dozza, 2014, 2015), as well as how their engagement affects their abilities and behaviors in different situations. This thesis focuses on the latter part. Such knowledge is important both in the development of vehicle systems and applications, such as human-machine interfaces and driver support systems, and in the creation of laws, regulations, and vehicle rating systems.

One of the ways vehicle manufacturers try to improve traffic safety is by developing driver support systems and, ultimately, systems for autonomous driving (AD). Importantly, for these systems to have the desired effect they must be not only safe, but also appreciated and used. It is expected that having the possibility to safely engage in cognitive tasks, such as work, while traveling will be a selling point for cars with AD systems in the future (Chuang et al., 2018). If more people use such systems, traffic safety can improve as more cars on the road drive autonomously—provided that AD is indeed safer than manual and assisted driving (although far from proven, several studies do indicate that this assertion is likely to be true; Bjorvatn et al., 2021; Elvik et al., 2020; M. M. Morando et al., 2018). The role of cognitive tasks in traffic safety could thus change from a safety concern to a safety enhancer by motivating the use of AD.

To understand (and counteract, if needed) the negative effects of cognitive tasks in both manual driving and AD, it must first be possible to measure and study such effects. A key difficulty is that reliable, validated measures for the continuous assessment of cognitive load (a mental response to cognitive task demand, described further in Section 3.1) are lacking.

Such measures would allow better generalizability of results, testing in less constricted and more realistic environments, and timely safety-system interventions based on driver state monitoring. There is a great interest in using physiological measures since they can provide continuous recordings of driver states without altering or interrupting the driving task (Longo, 2018). These measures can be used either together with other measures (such as self-reports and performance indices) to improve driver state assessments, or alone to assess driver states in situations where other measures are not sensitive or suitable (Lohani et al., 2019). But once again, results from different studies on the effects of cognitive load on physiological measures are inconclusive. Different measures respond to cognitive load differently in different studies (Tao et al., 2019). For physiological measures to be useful in a wide range of applications, the reason for this inconsistency needs to be understood.

In summary, car drivers'[1] engagement in cognitive tasks could affect drivers' behaviors and decisions in various ways in the different stages of a journey leading up to a potential crash (conceptually illustrated in Figure 1 below); in the pre-trip stage when decisions are made about how to transport oneself, in the routine (i.e., non-critical) driving stage, and in the critical event stage where driver (or vehicle) action is needed to prevent a crash. However, studies on the effects of cognitive task engagement on driver behaviors report apparently contradictory results, and there is a lack of studies on task engagement during AD in real traffic. In addition, measures for continuous and non-intrusive assessments of cognitive load are lacking, which limits the research and development of driver state-adapted systems.

In this thesis work, the role of cognitive tasks in traffic safety is therefore studied in the three stages mentioned above, and the possibility of measuring cognitive load using physiological measures is explored (see Figure 1).

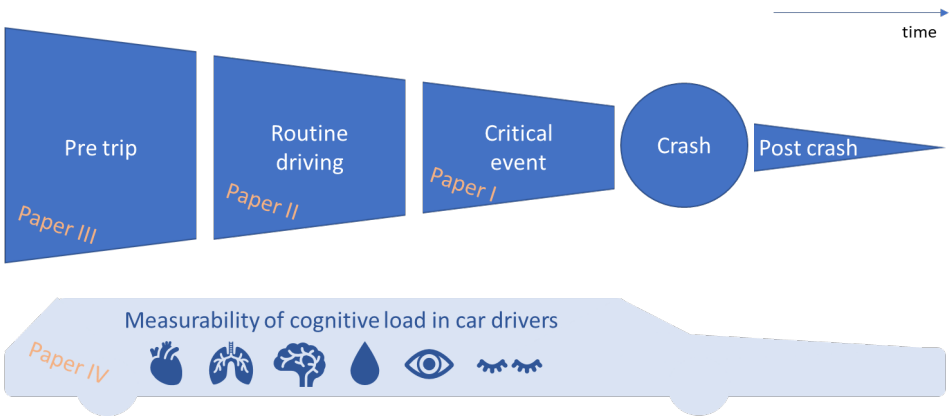

Figure 1. Illustration of how the papers included in this thesis relate to the stages of a journey which lead up to a potential crash (adapted from Eugensson et al., 2011) and to the measurability requirement spanning all stages.

[1] For the sake of simplicity, I will refer to the person in the driver seat as the driver, even during AD when the person is relieved from the task of driving.

2 Aim and objectives

This thesis aims to advance the understanding of task-induced cognitive load in the context of traffic safety, by accomplishing two objectives: A) to better understand how drivers' involvement in cognitive tasks can affect safety-relevant driver behaviors and B) to provide methodological guidance on how to assess cognitive load in car drivers using physiological measures. For this purpose, the cognitive load construct and the contextual influences on its effects will be considered.

The work in this thesis includes a high level of detail, decomposing observed behavioral and physiological responses to cognitive task engagement. Doing so provides a high time resolution of effects, an improved understanding of interactions between various effects, and a more detailed theoretical understanding of the effects of cognitive tasks on drivers' behaviors and physiological responses.

The main research questions addressed in this thesis are:

1. How do cognitive tasks affect drivers' response times in an unexpected safety-critical event? (Paper I; Objective A)
2. How do cognitive tasks affect drivers' glance behavior in routine driving? (Paper II; Objective A)
3. To what extent are drivers able to engage in non-driving-related tasks while traveling in AD mode in real traffic? (Paper III; Objective A)[2]
4. How can cognitive load in car drivers be assessed using physiological measures? (Paper IV; Objective B)

[2] Drivers' ability to engage in non-driving-related tasks during AD may influence their willingness to invest in and use cars with AD systems. In other words, the ability to engage in non-driving-related tasks may influence pre-trip decisions (namely, how to transport oneself), which in turn affect traffic safety (presuming AD will indeed increase traffic safety).

3 Background

3.1 Defining cognitive load

Clear definitions are important for comparisons between studies and for generalizations of study results. This chapter therefore reviews cognitive load definitions and typical cognitive load implementations in traffic research, as well as explains key terms used in this thesis.

Despite being rigorously studied in the last few decades, the concept cognitive load lacks both an agreed-upon definition and an agreed-upon name. Mental load, mental workload, and cognitive load are common terms that are used synonymously (Gwizdka, 2021). Moreover, the concept of load has been used both to describe the demands placed on the human and the impact on the human caused by these demands (Mehler, Reimer, & Zec, 2012). In this thesis, the two terms cognitive demand and cognitive load will be used to describe the demand itself and the impact it has on the human, respectively. Thus, a cognitive demand is something that results in cognitive load.

Van Acker et al. (2018) performed a comprehensive concept analysis on cognitive load, concluding that the various definitions all agree that cognitive load is the spending of cognitive resources of finite capacity to meet cognitive (task) demands. As an example, an early (and still-used) definition from O'Donnell and Eggemeier (1986) states that cognitive load "refers to that portion of the operator's limited [resources] actually required to perform a particular task" (p. 42-2). However, neither in their definition nor in many others is it clear what these (cognitive) resources are (Cohen, 2017). O'Donnell and Eggemeier (1986) simply refer to them as processing facilities that enable task performance. Van Acker et al. (2018) state that spending them entails task-directed attention and effort through working memory functions. Others describe them as neural mechanisms (Engström, Monk, et al., 2013) or attentional resources (M. S. Young et al., 2015) that enable cognitive control.

Cognitive control is a concept closely related to cognitive load. It can be described as the human ability to override habitual and automatic behaviors in favor of goal-directed behaviors and is often regarded as synonymous with executive function (Cohen, 2017). When subjects exert cognitive control, activity increases in the brain's frontal cortex (Botvinick & Cohen, 2014), as well as in other brain areas (Barkley, 2012). Multiple cognitive functions are involved in achieving cognitive control; indeed, more than 30 different functions have been suggested (Eslinger, 1996)—including attention, working memory, error monitoring, inhibitory control, and planning (Helfrich & Knight, 2016). Further, cortical arousal is a prerequisite for (Sadaghiani & Kleinschmidt, 2016) and modulator of (Grueschow et al., 2020) cognitive control. However, cognitive control, like many of its underlying cognitive functions, suffers the same problem as cognitive load does; there is no agreed-upon definition (Barkley, 2012). There is also a lot that is not known about the neurocognitive processes that enable cognitive control and many other cognitive functions (Barkley, 2012).

The lack of neurocognitive knowledge and precise definitions make the relationships between the different functions unclear (Cohen, 2017), just as it is unclear which functions should be

included in the cognitive control and cognitive load constructs. For example, Van Acker et al. (2018) consider functions used for stimuli detection and response execution, such as perception and motor control, as cognitive load functions, while Engström, Monk, et al. (2013) do not.

Further, Van Acker et al. (2018) argues that emotional load (i.e., emotional responses to task demand) should be differentiated from cognitive load. As they point out, it is implied that cognitive load gives rise to a subjective experience, since self-ratings are frequently employed in cognitive load assessments (e.g., NASA Task Load Index; Hart & Staveland, 1988); however, the relations between the subjective experience, emotional response, and cognitive load are very ambiguous. They argue that because emotions are not necessary to perform cognitive tasks, they should not be considered parts of the cognitive load construct. Typically, cognitive load definitions do not specifically mention emotions. However, emotions have a significant impact on human behaviors, cognitive processes, and task performance (Barkley, 2012; Tyng et al., 2017). For example, stimuli with an emotional component capture attention and trigger faster responses than neutral stimuli, emotional events augment memory, and decisions are affected by previous emotional responses (Dolan, 2002). Kluwe-Schiavon et al. (2017) even suggest in their dynamic executive function hypothesis that the role of executive function (which they, unlike many others, do not consider to be synonymous with cognitive control) is to balance the level of activity in the brain's salient network (related to emotions and involving sub-cortical structures) and the executive control network (related to cognitive control and involving frontal-cortical structures) for optimal performance. In particular, emotional stress is considered a motivator for problem-solving, while cognitive control is necessary to solve the problem; appropriate activity in both systems is necessary for situation-appropriate behaviors (Kluwe-Schiavon et al., 2017).

Regardless of exactly which functions are included in the cognitive load construct, it is clearly a multidimensional construct that can take many forms (Matthews et al., 2015). That is, cognitive load can include different cognitive functions (including emotions) at different times.[3] The type and amount of cognitive load is determined by the specific task demands and by the environmental context in which the task is performed, as well as by the characteristics of the human performing the task (M. S. Young et al., 2015).

Humans self-regulate their cognitive load by modifying their strategies and actions to meet task goals (e.g., to drive to a certain location, or to perform a cognitive task when being asked to do so) as well as personal goals (e.g., to enjoy the drive, or to make the test leaders happy by doing really well in the cognitive task) (Hancock & Matthews, 2019; Kircher & Ahlstrom, 2017). Their behavior, and the level and type of cognitive load, is thus the result off a compromise between many different goals, including the goal of not investing more effort than is deemed worth the cost (Cohen, 2017). Further, mental states such as emotions, sleepiness, and fatigue also have an impact on both task performance and the type and level

[3] The ISO standard for the detection-response task (DRT; International Organization for Standardization [ISO], 2016) (which use the same definition of cognitive load as Engström, Monk, et al., 2013) does, however, argue that cognitive resources can be viewed as a single resource because cognitive tasks of different kinds all activate mostly the same areas in the brain's frontal cortex (in particular the lateral prefrontal cortex and the anterior cingulate cortex). This "smallest common denominator approach" to defining cognitive load thus excludes any mental responses that are not always (or usually) involved in cognitive control.

of cognitive load by affecting brain functioning and, consequently, strategies and actions. For example, anxiety tends to lead to greater intolerance of uncertainty and can cause perseverative information seeking and impaired decision making (Wever et al., 2015), positive emotions unrelated to the task increase distractibility (Goschke & Bolte, 2014), task-induced fatigue leads to sleep-related brain activity that worsens performance (Chong & Baldwin, 2021), psychosocial stress impairs working memory performance (Schoofs et al., 2008), and negative emotional stress delays and reduces task-related brain responses in the prefrontal cortex (Gärtner et al., 2015).

Interestingly, while large parts of the scientific community seem to agree that human cognitive performance is restricted by limited resources, some voices question this assumption (e.g., Chong & Baldwin, 2021; Kurzban, 2010). Perhaps the most clear-cut objection is that limited resources alone cannot explain human performance degradation. For example, while prolonged task execution causes fatigue and performance degradation (Kato et al., 2009), performance can be restored with reward manipulations (Hopstaken et al., 2015), which contradicts the assumption that performance degrades solely due to the depletion of resources (as suggested by Baumeister et al., 1998). Also, the concept of limited resources alone cannot explain observed performance degradations in non-challenging tasks, where plenty of resources should be available. It has been suggested that underload causes performance degradations because the pool of cognitive resources shrinks during low task demand (M. S. Young & Stanton, 2002). However, Dehais et al. (2020) argue that such high-level theories are hard to operationalize, and that performance degradations seen when cognitive demand is either too high or too low instead need to be understood from a neuroergonomic perspective: that is, with respect to the neurobiological mechanisms that underlie performance. The development of advanced neuroimaging techniques has enabled great progress in this field in the last few decades (Dehais et al., 2020). More is thus becoming known about how mental states affect brain functions (and, consequently, performance) through changes in, for example, levels of neurochemicals, interactions between brain networks, and neuronal firing rates (Chong & Baldwin, 2021; Dehais et al., 2020; Grueschow et al., 2020; Shine et al., 2016).

Highly relevant when discussing cognitive load is humans' poor ability to multitask. C. D. Wickens' (2002, 2008) multiple resource theory, developed to predict dual task performance, is well known to human factors practitioners. It suggests that humans have multiple types of cognitive resources, each with a limited capacity, and that we perform better at simultaneous tasks which are different from each other, since they draw on different resources, than at those which are similar and thus draw from the same resource pool.

From a neurological perspective, it has been suggested that humans' limited ability to multitask with similar tasks is not due to depletion of a certain category of resources, but that the very purpose of cognitive control could be to prevent inappropriate multitasking (Botvinick & Cohen, 2014; Meyer & Kieras, 1997) in order to prevent errors caused by cross-talk (i.e., interference) between closely located and simultaneously active neural assemblies (Cohen, 2017). The brain prevents cross-talk and enhances focused attention by means of inhibitory mechanisms which dampen neural activities that are task- or stimulus-irrelevant in order to enhance the signal-to-noise ratio for the task-specific neural activity (Levy & Anderson, 2002). These mechanisms involve lateral inhibition, in which active neurons dampen the activity of their neighboring neurons (Burke et al., 2017), and inhibitory

control, in which unprioritized neural activity is inhibited by higher-level cortical areas (i.e., through top-down control; Munakata et al., 2011). While neural inhibition can improve task performance, excessive inhibition can also explain some negative effects observed during high levels of task demand and stress (Dehais et al., 2019). For example, inhibition can keep the brain from considering new information and cause an individual to persist with an erroneous strategy (Dehais et al., 2020). An example is seen in overloaded aircraft pilots who, at all costs, attempt to land their craft in bad weather conditions, despite having the (obviously better) option of landing at another site with better conditions (Dehais et al., 2010).

In summary, cognitive load is usually defined as a spending of cognitive resources. These resources are considered to underlie various cognitive functions and to be of limited amount or capacity, which limits our cognitive abilities. It is, however, not clear what these resources actually are. Moreover, the idea of limited resources fails to explain some of the performance degradations seen during both high and low levels of cognitive demand. We therefore still lack a fully valid, agreed-upon definition of cognitive load.

For the purpose of this thesis, it is more important to acknowledge that cognitive load is a multidimensional construct than to agree on a precise definition of cognitive load. It is multidimensional because it consists of multiple mental responses, shaped not only by the cognitive task demand, but also by contextual and individual factors. Acknowledging that there are multiple, separate mental responses to cognitive task demands is important—both for understanding the mechanisms underlying changes in driver behaviors and for assessing cognitive load using physiological measures.

3.1.1 Operationalization of cognitive load in driving research

While the discussion above clearly shows that task-induced cognitive load is a multidimensional construct, driving studies interested in its effects on driver behaviors or physiological responses usually treat it as unidimensional. That is, when some cognitive demand is varied, the corresponding cognitive load is simply described as either increasing or decreasing. While convenient, this simplification has several shortcomings. To better understand what these are, this section provides a brief overview of the types of studies that are typically conducted.

To begin with, drivers' engagement in non-driving-related tasks (NDRTs) is often studied in terms of distracted driving (i.e., drivers directing their attention away from activities critical for safe driving and towards a competing activity; Streff & Spradlin, 2000). Driver distraction is further typically divided into cognitive, visual, and manual distraction, referring to drivers taking their mind off the road, their eyes off the road, and their hands off the steering wheel, respectively (Ahlström et al., 2012). Although most naturalistic tasks induce some degree of cognitive load/cognitive distraction (Engström et al., 2017), the term cognitive distraction is most often used to describe the impact from tasks that do not require drivers to take their eyes off the road or their hands off the steering wheel, but that require their cognitive engagement—that is, the impact from what is refered to as *cognitive tasks* in this thesis. I do not use the term *cognitive distraction* in this thesis because it typically implies a degraded primary task (i.e., driving) performance (Regan et al., 2011); an assumption I do not wish to make.

In experimental studies, different cognitive tasks are employed to induce cognitive load and are systematically varied. It can be, for example, artificial working-memory loading tasks or mental arithmetic tasks, or more lifelike tasks such as cell phone conversations or voice-controlled navigation. Other factors which may influence cognitive load, such as traffic density (Di Flumeri et al., 2018) and road geometry (Bruyas & Dumont, 2013), are either kept constant or varied systematically and included in the analysis (Faure et al., 2016). In contrast, in naturalistic driving studies, cognitive load cannot be experimentally induced. Instead, scenarios which include observable engagement in tasks which are assumed to induce cognitive load, such as cell phone conversations or conversations with passengers, are compared against similar baseline scenarios without the tasks (Dingus et al., 2019; Hickman et al., 2010; Victor et al., 2015). From this large variety of tasks and contexts, many different mental responses can be expected, with the common denominator that they all require some level of cognitive control. Because cognitive load is operationalized this way, it is (implicitly or explicitly) assumed that it is the need for cognitive control that affects the behaviors and physiological responses.

Notably, since cognitive tasks give rise to multiple mental responses, and since the definitions of these mental responses are often imprecise, overlapping, and not agreed upon (Cohen, 2017; Westbrook & Braver, 2015), the same tasks are sometimes used in different studies to explore different constructs. For example, the n-back task is a working memory loading task frequently used to study the effects of cognitive load (described further in Section 4.2 and employed in, e.g., Heine et al., 2017; Mehler et al., 2009; Medeiros-Ward et al., 2014). The n-back task has, however, also been used to study the effects of stress (Mathissen et al., 2021; Patel et al., 2018) and mental effort (Brouwer et al., 2014; Luong et al., 2019). Effects observed when performing the same cognitive task (effects which should consequently be comparable) are thus attributed to different mental responses (different constructs) in different studies, and it is not clear which mental response(s) actually gives rise to the observed effects.

A few driving studies have looked not only at the effects of varying levels of cognitive load but also at the effects of various types of cognitive load by using different types of cognitive tasks (e.g., Y. Wang et al., 2014; Zheng et al., 2005). Recarte and Nunes (2000), for example, found that verbal tasks (repeating words) and spatial-imagery tasks (mental rotation of imagined letters) had different effects on drivers' glance behaviors. (Note that there are plenty of studies exploring the differences between visually and non-visually demanding tasks, but only non-visually demanding tasks are the focus of this thesis.)

In summary, driving studies investigating the effects of cognitive load on driver behaviors and physiological responses either observe, or systematically vary, the level of cognitive load, using a variety of tasks requiring cognitive control. Typically, observed effects are (implicitly or explicitly) attributed exclusively to the changes in cognitive control. The influence of other potential mental responses is thus often overlooked, although it is well-known that cognitive tasks give rise to multiple mental responses which are shaped not only by the task but by the situation as a whole.

3.1.2 Nomenclature used in this thesis

The following section describes relevant terms used in this thesis and their relationships to each other (see Figure 2 and Table 1). Others may use the terms in slightly different ways; the descriptions provided here are not meant to be universal. The intent is to make it clear to the reader what is meant by the terms used. All descriptions build on the key assumption that the focus is on human subjects who are engaged in some non-driving-related task while driving.

To begin with, all tasks occur in a context which affects task performance via mental responses and behaviors. The context consists of factors that influence task demand (e.g., task difficulty, time pressure, type of perceptual demand), as well as of environmental factors (e.g., traffic density, type of vehicle, visibility) and of individual factors (e.g., personality, experience, level of arousal). The brain's *mental responses* are shaped by these contextual factors. The mental responses are both emotional responses and the activation/deactivation of cognitive functions and processes, such as attention, cortical arousal, perception, and motor control. Perception is the link between the brain and the environment, and motor control is the link between the brain and the subject's behavior. The mental responses that enable or facilitate cognitive control are called *cognitive load components*, and together they constitute the *cognitive load*. Cognitive load is thus an umbrella construct for multiple mental responses (or cognitive load components) that together enable or facilitate cognitive control in a specific situation. Mental responses that do not enable or facilitate cognitive control in the situation at hand are simply referred to as *other mental responses*. A certain mental response can thus be a cognitive load component at one time but not another. Moreover, it is not necessarily easy to determine whether a specific mental response is contributing to cognitive control and thus should be labelled a cognitive load components. Consider emotional stress, for example. Emotional stress may motivate effort investment and thus enhance task performance (Kluwe-Schiavon et al., 2017)—in which case it should be considered a cognitive load component. But if the stress becomes too high, it has detrimental effects on working memory and task performance (Schoofs et al., 2008)—in which case it can no longer be considered a cognitive load component, since it inhibits rather than enables cognitive control. Therefore, cognitive load components and other mental responses are typically not differentiated in this thesis; the overarching term 'mental responses' is used instead.

The mental responses together shape behaviors such as steering wheel movements, task responses, and gaze direction, both indirectly through motivation and effort (for example) and directly through motor control. These behaviors may affect both the mental responses (Leisman et al., 2016) and the context (e.g., a change in time headway, vehicle direction, level of fatigue), which in turn gives rise to new mental responses, and so on.

Of great importance for this thesis is the fact that the context, the mental responses, and the behaviors all give rise to physiological responses (e.g., changes in heart rate, pupil diameter, and skin conductance). The physiological responses can in turn also affect the mental responses and task performance. For example, physiological responses are important elements of emotions (Pace-Schott et al., 2019), and physiological arousal improves both memory encoding and memory recall (Critchley et al., 2013). The relations between mental and physiological responses are reviewed further in Sections 3.5.1 and 3.5.2.

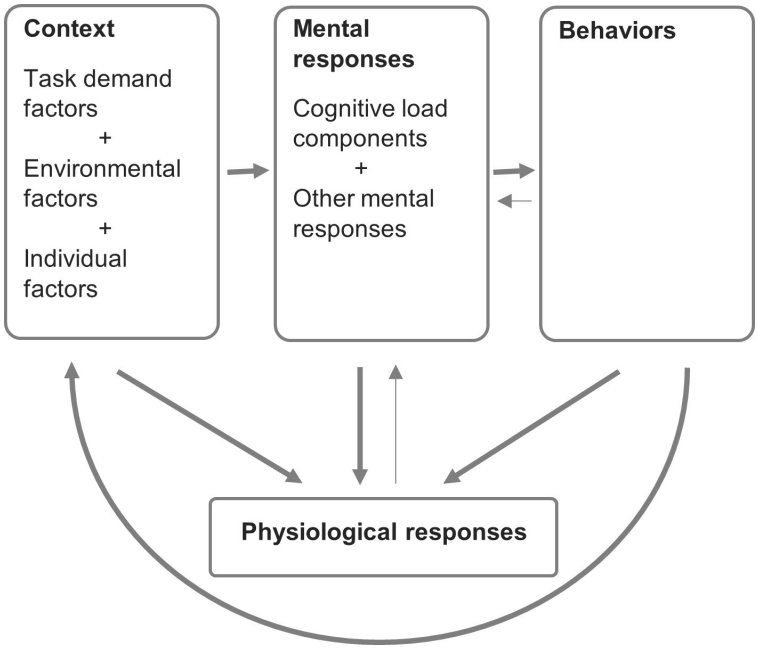

Figure 2. Relationships between contextual factors, mental responses, behaviors, and physiological responses. Thinner arrows indicate generally smaller effects, of less interest in this thesis.

Table 1. Compilation of key terms as defined and used in this thesis.

Term	Description
Mental response	The brain's response to stimuli or demands. It includes emotional responses, thoughts, and activation/deactivation of cognitive functions or processes (such as attention, cortical arousal, perception, and motor control).
Cognitive control	The ability to override habitual and automatic behaviors in favor of goal-directed behaviors (Cohen, 2017). Cognitive control is achieved through multiple cognitive functions (e.g., attention, working memory, and inhibitory control).
Cognitive load components	Mental responses that enable or facilitate cognitive control at a given time. This includes cognitive control functions which are needed for the task at hand, support functions such as cortical arousal, and emotional responses that improve cognitive control. A certain mental response can thus be a cognitive load component at one time but not at another.
Cognitive load	A construct used to collectively refer to the cognitive load components: the collective mental response that enables or facilitates cognitive control. There may also exist mental responses to stimulus or demand that do not enable or facilitate cognitive control. These are referred to as other mental responses and do not form part of the cognitive load construct.
Cognitive task	Non-visual, non-driving-related task which requires cognitive control.*
Cognitive demand	The demands that a task (in this thesis, primarily cognitive tasks) places on cognitive functions for its performance. These demands depend on the task's characteristics (e.g., stimulus discriminability, stimulus presentation rate, and task complexity).
Task demand	All the demands a task places on a performer. It can include cognitive demand as well as other task demands, such as physical demands. Note that the task demands are task-specific, but the effects the task has on the performer in terms of mental responses are not; they depend on environmental and individual factors as well.
Driver	The person who drives the vehicle—or, during AD, the person in the driver's seat, who is then relieved from the task of driving.
Driving task	The task of controlling or operating the vehicle, including supervising vehicle systems and the traffic environment during assisted driving. Strategic choices, such as which route to take, are not referred to as driving tasks in this thesis.

*Recall that the term *cognitive task* is used instead of the acronym NVCLNDRT (Non-Visual, Cognitively Loading, Non-Driving-Related Task). That is not to suggest that tasks that are driving-related or visually demanding cannot also require cognitive control (indeed, they most often do).

The following sections will review the effects of cognitive tasks on driver behaviors.

3.2 Effects of cognitive tasks on response times (critical situations)

In a safety-critical event, a fast and appropriate driver response is often crucial to avoid a crash. However, there is a great concern that cognitive load causes increased response times, resulting in an increased risk of crashes in critical situations (Merat & Jamson, 2008; Patten et al., 2004; Strayer & Fisher, 2016). The concern stems from the large number of experiments that have shown increased response times to different stimuli, such as artificial detection tasks (Bruyas & Dumont, 2013; Conti et al., 2012; Engström, Larsson, et al., 2013; Mantzke & Keinath, 2015) and braking lead vehicles (Alm & Nilsson, 1995; Engström et al., 2010; Salvucci & Beltowska, 2008), when drivers were performing cognitive tasks. However, naturalistic data has, as previously mentioned, actually shown a decreased risk of rear-end collisions during cell phone conversations (e.g., Victor et al., 2015), which looks at first glance like a contradictory finding.

Interestingly, a meta-study by Engström (2010) showed that the experimental design itself greatly influences the effect size of cognitive tasks on brake response times in lead vehicle braking scenarios: the effect size was found to be dependent on the time headway (THW) to the lead vehicle at the moment it started braking—that is, on the scenario criticality. In studies where that initial THW was large, cognitive tasks had a large effect on the response time, while in studies where the THW was small, cognitive tasks had a much smaller effect on the response time. In line with this finding, response times in rear-end crashes and near-crashes culled from naturalistic data also depend on the scenario criticality (Markkula et al., 2016).

A theoretical framework called the cognitive control hypothesis, formulated by Engström et al. (2017), seems helpful for reconciling the seemingly inconsistent effects of cognitive tasks on response times in different tasks and in scenarios with varying criticality.

3.2.1 The cognitive control hypothesis

The cognitive control hypothesis suggests that cognitive load selectively impairs driving subtasks that rely on cognitive control but leaves automatic behaviors unaffected (Engström et al., 2017). Automatic behaviors are effortless and don't require active control or attention by the subject, while the opposite is true for cognitively controlled behaviors (Schneider & Shiffrin, 1977). Engström et al. (2017) propose that the Guided Activation Theory (GAT) can explain the relationship between controlled and automatic behaviors and how automaticity develops. The GAT describes in neurobiological terms how all behaviors lie somewhere on a continuum between controlled and automatic; the degree of automaticity is determined by neural-pathway strength (Botvinick & Cohen, 2014; Cohen et al., 1990; E. K. Miller & Cohen, 2001). Successful employment of neural pathways increases their strength, leading to a higher degree of automaticity (Cohen et al., 1990). Consequently, tasks with a consistent stimulus-response mapping (i.e., tasks in which a certain stimulus always leads to the same response) can become more automatized through extensive practice.

To successfully deal with novel tasks and tasks in which a certain stimulus can require different responses, humans are capable of flexible and non-routine behaviors in which stronger pathways (i.e., more automatic responses) are overridden in order to activate weaker pathways that are more relevant. This capability is achieved through the employment of

cognitive control (Botvinick & Cohen, 2014; E. K. Miller & Cohen, 2001). If a behavior with weaker pathway strength is not activated through cognitive control, a behavior with stronger pathways will instead activate, resulting in routine, or automatic, behavior.

In the driving context, this means that when drivers encounter new or complex situations, such as when driving through a busy, non-signalized intersection, they need to apply cognitive control in order to adapt their behavior to fit the current situation. How much cognitive control is needed depends both on the situation's complexity and the driver's experience in similar situations. However, applying cognitive control requires effort, and drivers will only do so to the extent they believe it is worth the cost. That is, their behavior is typically satisficing, rather than optimizing (Kircher & Ahlstrom, 2017; Summala, 2007). Also, because humans' ability to multitask is limited (as discussed earlier), when drivers engage in cognitive tasks (which require cognitive control) they are less capable of applying cognitive control to driving, so the driving comes to depend more on automatized behaviors (Engström, Victor, et al., 2013; Engström et al., 2017).

The previously mentioned response-time studies can be viewed from the cognitive control hypothesis perspective by considering whether the stimulus to which the driver is to respond elicits automatized responses. Tasks that are novel to the participants are not expected to elicit automatic responses due to insufficient stimulus-response mapping, and the hypothesis thus predicts increased response times for cognitive tasks in those cases. One such task which is frequently used is the ISO standardized detection-response task, DRT (ISO, 2016), which requires participants to press a button as fast as possible upon detection of a visual or tactile stimulus presented at irregular intervals of 3–5 s. Although the task is practiced before trials, it is not practiced to the extent that full automaticity can be expected (ISO, 2016; Shiffrin & Schneider, 1977). In line with the hypothesis' prediction, numerous studies have shown increased response times to the DRT stimulus during cognitive task execution compared to a baseline (no task) condition (Bruyas & Dumont, 2013; Conti et al., 2012; Engström, Larsson et al., 2013; Mantzke & Keinath, 2015; Merat & Jamson, 2008; Patten et al., 2004).

In lead vehicle braking studies, it is necessary to take into account what the drivers are responding to (what the stimulus is) in order to assess the degree of automaticity. Potential stimuli are brake lights and visual looming of the lead vehicle (the optical expansion of the lead vehicle on the driver's retina). A looming object on a collision course with the subject unconsciously captures attention (Lin et al., 2008) and elicits automatic avoidance responses (Náñez, 1988). The cognitive control hypothesis thus predicts that engagement in cognitive tasks would not interfere with brake responses to a looming lead vehicle. Brake lights, on the other hand, are not likely to elicit automatic responses, because they are often encountered in situations which do not require an immediate brake response and are thus not consistently mapped to a brake response. Therefore, the cognitive control hypothesis predicts that cognitive tasks will interfere with responses to brake light onset (causing longer response times), provided that visible looming is absent.

Cognitive control can also be employed to improve response times in studies where lead vehicle braking is anticipated—either due to instructions, cues in the traffic environment, or scenario repetition. Response times in studies with such designs are hence also expected to be affected by cognitive tasks.

Results from lead vehicle braking studies are in line with these predictions. The vast majority of the studies have employed study designs which allow for or require cognitive control (Alm & Nilsson, 1995; Engström et al., 2010; Lee et al., 2001; Salvucci & Beltowska, 2008; Strayer et al., 2013). An exception is the work of Muttart et al. (2007), who studied response times in a lead vehicle braking scenario where brake lights were de-activated and the slowing of the lead vehicle could either be predicted by downstream traffic events or occurred for no apparent reason. In line with the cognitive control hypothesis, the authors found no effect of cognitive task execution on the response distance in the un-cued (looming only) condition, whereas in the cued condition the cognitive task significantly delayed the response. Similarly, Baumann et al. (2008) found that cognitive task execution delayed driver responses when there was a looming obstacle after a curve that had been cued by a warning sign, but not when the sign was absent and the response was solely triggered by the looming obstacle.

In sum, the seemingly inconsistent effects of cognitive tasks on response times appears to be well explained by the cognitive control hypothesis (Engström et al., 2017). That is, response times are increased in tasks and events in which cognitive control is employed when the response is generated, such as those that require anticipation or less practiced behaviors. On the other hand, responses which are automatically triggered (e.g., by visual looming) are not prolonged. However, because very few studies have employed such study designs, further research is needed to confirm these findings. Therefore, the effects of cognitive tasks on response times in an unexpected event with visual looming were studied in Paper I.

3.3 Effects of cognitive tasks on glance behaviors (routine driving)

Car driving under normal conditions requires continuous adaptation to the traffic environment based on visual information. To ensure appropriate behavior and safe driving, it is thus necessary that drivers maintain a glance behavior appropriate for the current traffic situation, where areas most likely to have important information are appropriately scanned.

The most common gaze direction for car drivers is towards the future path (Ahlström et al., 2012; Harbluk et al., 2007; Victor et al., 2005; Y. Wang et al., 2014). Looking at the future path is important for lane keeping and path planning (Land, 2006), as well as for noticing objects on an immediate collision course (Lehtonen et al., 2012). When drivers perform cognitive tasks, they increase the time spent looking towards the future path (Harbluk et al., 2007; Reimer et al., 2013; Y. Wang et al., 2014) and exhibit a reduced variability in gaze direction (Recarte & Nunes, 2000, 2003; Reimer et al., 2012; Reimer, 2009; Victor et al., 2005; Y. Wang et al., 2014). This altered glance behavior is referred to as the gaze concentration effect. It has been suggested that gaze concentration is a compensatory visual attention strategy that drivers use to cope with the increased cognitive load by prioritizing the most safety-relevant area: the road ahead (Recarte & Nunes, 2000, 2003). Others have, however, argued that gaze concentration arises because drivers' cognitive resources are depleted, making the drivers less capable of situationally adapted gaze control such as scanning for potential hazards (Biondi et al., 2015).

Most studies of drivers' visual behavior are done on highways or main roads where the visual behavior is studied over extended periods of time (Recarte & Nunes, 2000, 2003; Reimer et al., 2012; Reimer, 2009; Victor et al., 2005; Y. Wang et al., 2014). These settings do not

require many off-path glances (i.e., glances to locations other than the future path); increased time spent glancing towards the future path could therefore be beneficial for safety. Indeed, the gaze concentration effect has been suggested as an explanation for the finding that drivers conversing on cell phones have a decreased risk of rear-end collisions (Dingus et al., 2019; Victor et al., 2015).

In contrast, a different visual scanning behavior is needed in more complex traffic environments, such as intersections and cities, where there is more safety-relevant information in off-path locations. Glances to specific off-path locations are then more important for threat detection and situation awareness. However, even studies in such environments have found increased gaze concentration during cognitive task engagement, with fewer glances towards situationally relevant places other than the future path. For example, when drivers are engaged in cognitive tasks, they are less likely to look in rear-view mirrors before lane changes (Muttart et al., 2007) and to glance towards traffic lights (Harbluk et al., 2007) and the right and left at intersections and crosswalks (Biondi et al., 2015; Strayer et al., 2015). This deteriorated scanning of safety-relevant areas in complex traffic environments is likely to be detrimental to traffic safety (Harbluk et al., 2007).

In summary, car drivers' engagement in cognitive tasks induces gaze concentration: the drivers look more towards the road ahead and have fewer glances towards off-path locations. The effect seems to be independent of the traffic environment and the amount of relevant information in off-path locations. But because most studies to date have only analyzed these effects in aggregated data (i.e., over larger driving segments), there is a lack of knowledge about the gaze concentration effect's relation to environmental and cognitive demands. Paper II addresses this knowledge gap by exploring drivers' gaze behavior in relation to variations in cognitive demand and situationally relevant objects in off-path locations.

3.4 Cognitive task engagement in assisted and autonomous driving

Driver assistance systems are being developed to improve traffic safety by reducing the number of accidents caused by driver errors (Cabrall et al., 2019). An increasing number of vehicles on the market are equipped with systems that can warn the driver of safety-relevant events, such as forward collision warning systems and blind spot warning systems (Highway Loss Data Institute [HLDI], 2020). Some systems can also act if the driver fails to take action—for example, automatic emergency braking (HLDI, 2020). Many vehicles also have systems for assisted driving which can alleviate the driving task demand by providing longitudinal and/or lateral regulation support (Cabrall et al., 2019). Importantly, these systems cannot handle all traffic situations and continuous regulation is neither guaranteed nor intended. The driver therefore always remains responsible for overall vehicle control when using these systems and must intervene if the systems' actions do not suit the traffic situation (Cabrall et al., 2019).

At the same time, systems for AD are also being developed (M. Wood et al., 2019). The main difference from a user perspective is that when AD is engaged, the driver is no longer responsible for overall vehicle control and does not need to supervise the driving (United Nations Economic Commission for Europe [UNECE], 2020). Descriptions of AD systems typically refer to the set of conditions in which they can operate (the Operational Design

Domain; ODD) and whether a human in the vehicle might be requested to take over manual driving control at some point. Several nomenclatures for describing different types of AD systems and their capabilities exist. For example, SAE International has in their SAE J3016 automation taxonomy (SAE International, 2021) classified six levels of automation from Level 0 (manual driving) to Level 5 (full automation). For the purposes of this thesis, such a detailed nomenclature is not necessary. Here, it is sufficient to have one key distinction: systems for AD do not require a driver to supervise the driving while the system is activated, while systems for assisted driving do.

It follows that the way drivers engage in NDRTs, as well as the effects of their task engagement on traffic safety, are likely to differ between manual driving, assisted driving, and AD. It has been shown in naturalistic studies that drivers who are experienced with assisted driving systems increase the amount of time they engage in NDRTs when the system is activated compared to during manual driving (Dunn et al., 2019; Malta et al., 2012). This propensity may make them less prepared to take over control if the assistance stops regulating (Dunn et al., 2019), as indicated by increased response times to unexpected critical events in simulator studies (de Winter et al., 2014). Interestingly in naturalistic data, it has been observed that when using systems for assisted driving, drivers direct their gaze towards the future path before the system starts braking, indicating that they anticipated the critical situations (Tivesten et al., 2015). Monotonous tasks, such as supervising automation, may also lead to drowsiness (Matthews et al., 2019), which in turn can degrade driver performance and traffic safety (Dingus et al., 2016). In such cases, cognitive tasks, such as cell phone conversations, may mitigate driver drowsiness and actually improve traffic safety (Wijayaratna et al., 2019).

When using AD systems, drivers' engagement in NDRTs should not pose any problem since the drivers are not required to supervise. Instead, the AD system shall perform the driving task and manage all situations (including system failures) without exposing the vehicle occupants or other road users to any unreasonable risks (UNECE, 2020, paragraph 5.1.1). The assumption is thus that the system provides more safety than a human driver would (M. Wood et al., 2019). However, drivers' activities and mental states could impact their ability to take over control when the vehicle exits its ODD (Weaver & DeLucia, 2020; Sato et al., 2020). When drivers no longer supervise or participate in the driving, their situation awareness could deteriorate and make them unprepared to take over control in the case of a take-over request (TOR; de Winter et al., 2014). Engagement in NDRTs could potentially further impair their situation awareness as attention is focused on the NDRT (Marti et al., 2022). In addition, when a TOR is issued while the driver is engaged in an NDRT, task-switching in terms of re-directing attention from the NDRT to the driving task is necessary—which may also lead to more errors and slower responses (Monsell, 2003). Indeed, studies have found deteriorated take-over performance during engagement in NDRTs, especially when the NDRTs impose a visual or motoric demand (see Weaver & DeLucia, 2020, for a review). Moreover, as with assisted driving (but potentially even more critically), driver drowsiness might increase when the driver is relieved from driving (Jarosch et al., 2019; Sato et al., 2020) and may also impair take-over performance (Vogelpohl et al., 2019). In that case, engagement in NDRTs could help keep the driver alert, limiting the impairment in take-over performance (D. Miller et al., 2015; Schömig et al., 2015).

Since drivers' mental state may impact their ability to take over control after AD, it is crucial for traffic safety that vehicles hand over control to the driver in a safe manner, with respect to the driver's mental state. Also, should the driver become unavailable when the vehicle is approaching the end of its ODD (e.g., because the driver is sleeping), the vehicle must be able to handle this and make a safe stop (UNECE, 2020). Therefore, monitoring of drivers' mental state during AD is of utmost importance. This capability, together with the overall expected improvement of safety with AD (Elvik et al., 2020; M. Wood et al., 2019), suggests that society indeed would benefit from a shift from manual driving to AD. Provided that this is the case, the utilization rate of the systems becomes decisive for their effect on traffic safety; naturally, a system that is not used provides no safety benefit. Here, NDRTs could play an important role since having the possibility to safely engage in NDRTs while traveling could be a key reason for using AD systems.

Because AD permit drivers to engage in tasks that are otherwise not allowed or possible to perform while driving, it is assumed that AD systems will allow drivers to be more productive (Chuang et al., 2018). That is, AD is expected to free up time for working or engaging in other NDRTs during time spent in, for example, traffic jams and on motorways. Supporting this expectation, surveys have found that drivers feel positively about engaging in NDRTs while using AD systems (Hecht et al., 2020; Pfleging et al., 2016), and that most are looking forward to the arrival of such systems (Winkeler et al., 2019). In contrast, other surveys have found that many drivers fear, or are hesitant to use, AD (Americal Automobile Association [AAA], 2019; König & Neumayr, 2017; Tennant et al., 2019). In fact, it has been argued that psychological issues, rather than technical ones, are likely to constitute the greatest obstacles to the wide-spread use of AD in the future (Shariff et al., 2017).

Studies in simulators and on test tracks show that drivers do indeed engage in NDRTs during AD (Broström et al., 2019; de Winter et al., 2014; Wandtner et al., 2018). However, there is a lack of empirical studies that investigate to what extent people are able to engage, and maintain engagement, in NDRTs while using AD systems in real traffic; letting go of the driving task in an environment where there is no physical risk is not the same as doing so in real traffic. Behaviors observed in simulators and on test tracks may thus not generalize to real life situations (de Winter et al., 2012). Also, peoples' expectations about systems with which they have no practical experience may not accurately predict future real-life usage (Lindgren et al., 2021). Thus, results from surveys of expected AD usage should be interpreted with some caution. Therefore, studies are needed to find out if drivers can engage in NDRTs during AD even in real traffic, or if they need some additional support to do so—so that systems and vehicles can be designed so that they will be appreciated and used once on the market.

3.5 Measures of cognitive load

A key challenge for research on the effects of cognitive tasks is the lack of reliable and validated measures for continuous assessments of cognitive load. Such measures are important for several reasons (which will be discussed further in Section 6.6), such as for supporting generalizations about study results and comparisons between studies, for exploring the effects of cognitive tasks in less constricted and more realistic environments,

and for enabling timely safety system interventions based on driver state monitoring. This section gives a brief overview of the available methods for measuring cognitive load, and the following section focuses more deeply on physiological measures.

Measures of cognitive load are typically divided into three categories: self-report measures, performance measures, and physiological measures (O'Donnell & Eggemeier, 1986). In self-report measures, participants report their perceived load on one (e.g., Rating Scale Mental Effort [RMSE]; Zijlstra, 1993) or multiple (e.g., NASA task load index; Hart & Staveland, 1988) scales. Self-reports are typically easy to administer and analyze and are popular among many researchers because they reflect the participants' experienced cognitive load (Longo, 2018). The human consciousness is, however, limited and variations in cognitive load may not always be perceived (Annett, 2002). Self-ratings also suffer from peak-end effects, meaning that participants' ratings are biased toward the most intense (peak) experience and the end of the experience (Hsu et al., 2018). This limitation can be problematic, since ratings are often assumed to reflect the average experience over a certain period of time—typically a few minutes (Hertzum, 2021).

Interviews, although not quantitative measurements, are related to self-ratings since they are also based on users' experiences and self-assessments. Interviews as qualitative tools for cognitive load assessments deserve mention, since they can be used to explore and understand research subjects' opinions, behaviors, and experiences (Rowley, 2012). They are especially useful in novel fields of research when researchers do not quite know what to expect, or when they aim to get a deeper understanding of a certain phenomenon (Jamshed, 2014). Interviews are frequently used in driver studies to explore user experiences concerning, for example, human-machine interface design (Rukonic et al., 2021) or trust in AD (Wintersberger et al., 2021). Although studies on the effects of cognitive task engagement during driving use interviews extremely rarely to assess or understand cognitive load (see Wiberg et al., 2015, for an exception), interviews—if used together with quantitative measures—can offer a more detailed understanding of the still-underexplored multidimensionality of cognitive load (Schoonenboom & Johnson, 2017). However, it is very time-consuming to prepare, conduct, and analyze interviews, and the risk of bias is large (Qu & Dumay, 2011).

Performance measures of cognitive load are based on the assumption that performance degrades when cognitive resources are depleted (ISO, 2016). Performance is measured either in the primary or the secondary (or tertiary) task of a dual (or triple) task design (M. S. Young et al., 2015). Primary-task measures in driving studies are directly related to vehicle handling, such as steering wheel reversal rate (SWRR), standard deviation of lateral position (SDLP), and standard deviation of velocity. Advantages of these measures are that they can be collected in real time during natural driving without interrupting or otherwise altering the task of driving (Solovey et al., 2014). Disadvantages include that they are also affected by factors such as road geometry and traffic environment (Kountouriotis & Merat, 2016), and that the possibility of using them is limited or absent in assisted driving and AD. In addition to being used for cognitive load classification (Kountouriotis et al., 2016; Loeches De La Fuente et al., 2019; Rahman et al., 2020), primary task performance measures taken during drivers' cognitive task engagement are often used to research the cognitive tasks' effects on driver behaviors and traffic safety (Abd Rahman et al., 2020; Collet et al., 2010; Desmet & Diependaele, 2019).

Primary-task performance measures can also be derived from less naturalistic driving, such as in studies where lane keeping is observed after participants are instructed to do their best to stay in the middle of the lane (Chan & Singhal, 2015; Salvucci & Beltowska, 2008), thus making lane keeping a performance-optimizing task, rather than a performance-satisficing task, as it normally is (Summala, 2007). One ISO-standardized primary-task performance measure for assessment of cognitive demand is the lane change test (LCT; ISO, 2010). In the LCT, the demand of a secondary task is assessed by measuring drivers' performance in a simulated driving task requiring lane changes following unpredictable signs. This type of unnatural primary-task performance measure may be useful when comparing levels of cognitive load from different cognitive tasks (Bruyas et al., 2008; K. L. Young et al., 2011); however, because of restrictions imposed by the driving task, their usefulness is otherwise quite limited. Also, since the driving task has been altered, primary task performance measures from these type of studies should not be assumed to reflect performance in real-life driving (Engström et al, 2017).

Secondary- (or tertiary-) task performance measures instead assess performance in the secondary (or tertiary) task, assuming that the performance in that task depends on the participant's spare cognitive capacity (the cognitive resources not occupied by the other task(s)) (de Waard, 1996; M. S. Young et al., 2015). A popular task used for cognitive load assessment is the DRT, described in Section 3.2.1. The DRT is used to measure cognitive load by assessing the effects of cognitive demand on driver attention (ISO, 2016). It has been extensively tested and found sensitive to variations in cognitive demand in various settings (Bruyas & Dumont, 2013; Merat et al., 2015; Merat & Jamson, 2008). However, its usefulness is limited, because it makes the driving situation as a whole less realistic and alters the driver's mental state by increasing the overall cognitive load (Biondi et al., 2021).

A problem with performance measures (or at least measures of satisficing behaviors) is that performance can be kept relatively constant during small or moderate increases in cognitive demand by increased effort (Hertzum, 2021), making such measures insensitive to smaller variations in cognitive load (Chen et al., 2016; de Waard, 1996; Mehler et al., 2009).

The third type of measurement, physiological measures, reflects bodily responses to cognitive demand. Many of the ways the body prepares to deal with, and then actually deals with, challenges are observable through physiological changes. Advantages of physiological measures are that many of them can be taken continuously, they do not alter or interrupt the driving task, and they are arguably not sensitive to bias since they mostly occur unconsciously. In addition, unlike primary task performance measures which are dependent on the act of driving, they can still be recorded during AD (Collet & Musicant, 2019). However, they are influenced by many different factors, such as emotions (Balters & Steinert, 2015), fatigue (Hu & Lodewijks, 2020), environmental conditions (Dawson et al., 2016), nicotine usage (Quintana & Heathers, 2014), and food digestion (Sauder et al., 2012). Also, not all measures can (to date) be recorded non-intrusively (i.e., without measurement equipment mounted on the body), and the precision of non-invasive recording techniques is still not as good as the more intrusive techniques (Arakawa, 2021; Cori et al., 2019; Wusk & Gabler, 2018). Physiological measures are the measurement category in focus in this thesis, and they will be reviewed further in the section below.

In assessments of cognitive load, it is of great importance to remember that cognitive load is a multidimensional construct, even though it often is assessed as if it is unidimensional (as discussed in Section 3.1.1 and Paper IV). In fact, analyses show that different measures affected by cognitive load do not necessarily correlate with each other (Matthews et al., 2015), whether they are from the same measurement category or not (Hancock & Matthews, 2019). The different measures thus appear to reflect different mental responses to cognitive demand. Consequently, it is possible to get a more detailed assessment of the multidimensional cognitive load using multiple measures, preferably from different measurement categories (de Waard, 1996; Hancock & Matthews, 2019).

3.5.1 Physiological responses to cognitive demand

This section provides an overview of physiological measures that can be used for cognitive load assessments as well as an understanding of *why* the different measures correlate with changes in cognitive demand. Such an understanding is important in order to recognize the usefulness of physiological measures for multidimensional cognitive load assessments (i.e., assessments of multiple mental responses to cognitive demand). Also, understanding the mechanisms underlying physiological responses makes it possible to use the measures to learn more about how cognitive tasks affect drivers' mental responses and behaviors. (For further reading, see Paper IV.) In the next section (3.5.2), some of the greatest challenges with physiological recordings and analyses (namely issues with signal artifacts and multiple relationships between physiological measures and mental states) are briefly presented.

To begin with, increased cognitive demand causes increased activity in the pre-frontal cortex and other areas of the brain that enable cognitive control and other task-specific functions (such as perception and motor control) (Duncan & Owen, 2000; Tomasi et al., 2007). These changes in brain activity can be observed with various methods, all with their pros and cons. Functional magnetic resonance imaging (fMRI) is a useful measurement technique for studying where in the brain changes in activity occur, thanks to its high spatial resolution and full-brain coverage (Scarapicchia et al., 2017). Participants are placed in a large scanner where changes in blood oxygenation in response to changes in neural activity are measured during task execution (Scarapicchia et al., 2017). fMRI studies are, however, very restricted because the participants must be completely still, most often lying down, during measurements.

Functional near-infrared spectroscopy (fNIRS) and electroencephalography (EEG) are methods for studying changes in brain activity in less restricted environments, such as in more realistic driving simulators and real cars (Baker et al., 2017). Similar to fMRI, fNIRS measures blood oxygen levels—but with a wearable device that allows more movement (Scarapicchia et al., 2017). The spatial resolution is lower than fMRI, and unlike fMRI, fNIRS can only measure activity in the brain's outer layer (the cortex). Also, although fNIRS allows for some motion, it is still prone to artifacts during motion, and the neural signals are contaminated with superficial scalp signals (Scarapicchia et al., 2017).

EEG is measured with electrodes on the scalp that record the electrical activity in the brain's outer cortex which derives from neuronal communication. EEG recordings have very high temporal resolution (at the ms level; Burle et al., 2015), but low spatial resolution in comparison to fMRI (although it can be improved with signal processing methods that

incorporate activity from multiple electrodes; Burle et al., 2015). Because the electrical signals from the brain are very weak (mostly below 50 microvolt; Daly et al., 2012), EEG recordings are very sensitive to artifacts, both from the environment and the body itself (e.g., eye and muscle activity; Tatum, 2014). Studies using EEG most commonly perform spectral analyses or look at oscillatory responses to discrete events by computing event-related potentials (ERPs). In spectral analyses, the signals' oscillatory waveforms are split into a number of frequency bands (typically delta: 1-4 Hz, theta: 4-8 Hz, alpha: 8-13 Hz, beta: 13-22 Hz, gamma: 22-80 Hz) and the power in the different bands is studied (Choi & Kim, 2018). Although it is still not fully understood *why* the brain exhibits activity in these frequency bands, more is known about how different mental states affect the activity in the different bands (particularly in laboratory environments). During increased cognitive demand, theta power generally increases over the frontal cortex (Ismail & Karwowski, 2020), which is suggested to reflect a need for cognitive control (Cavanagh & Frank, 2014; Cavanagh & Shackman, 2015). Further, alpha power generally increases in task-irrelevant sensory areas and decreases in task-relevant sensory areas (Sadaghiani & Kleinschmidt, 2016). Alpha activity was previously thought to represent an "idling" of the brain (Pfurtscheller et al., 1996), but theories today suggest it is an active functional mechanism that supports information processing and cognition. For example, Sadaghiani and Kleinschmidt (2016) hypothesize that alpha oscillations support purging of accumulated neural noise and task-irrelevant information in the brain at ~10 Hz in order to prepare the brain for incoming (e.g., sensory) signals. This hypothesis provides an explanation for some of the observed effects of different mental states on alpha power, both on a local and a global scale. Namely, global alpha power increases during sustained alertness because the brain needs to remain susceptible to sensory input. Local alpha power, on the other hand, decreases in task-specific areas during task engagement to improve selective attention and information processing by permitting prolonged accumulation of activity in these areas.

ERPs, in turn, can provide insight into cognitive operations related to sensory stimulus and behaviors (Woodman, 2010). Because single responses are small and easily masked in simultaneous, non-stimulus-related, brain activity, ERPs are typically derived by averaging many EEG signal segments centered around a repeated stimulus, such as an auditory tone or blinking light (Choi & Kim, 2018). Many times these stimuli are artificial elements that make the experiments quite unnatural and risk altering the driver's mental state and/or interfering with the driving task (Ahlström et al., 2020). One ERP measure that does not require an artificial stimulus and thus allows for a more ecologically valid study designs is the eye fixation-related potential (EFRP). As the name implies, the EFRP is time-locked to eye fixation onsets, occurring when a gaze shift ends and the inflow of new visual information starts (Takeda et al., 2012). Typically measured over occipital areas, The EFRP is thought to reflect the depth of visual information processing (Itoh et al., 2006).

As mentioned in Section 3.1, one of the fundamental mental responses to cognitive demand is cortical arousal, which is necessary for cognitive control (Sadaghiani & Kleinschmidt, 2016). One measure that is sensitive to variations in cortical arousal is pupil diameter (Eckstein et al., 2017), due to the fact that the locus coeruleus, one of the brain areas that controls pupil diameter, also controls the distribution of the arousal-promoting hormone norepinephrine in the brain (Joshi & Gold, 2020). Numerous studies have demonstrated increased pupil diameters during increased cognitive demand (see van der Wel & van Steenbergen, 2018, for

a review). Although most driving studies on pupil diameter are simulator studies (e.g., Hammel et al., 2002; Niezgoda et al., 2015), some have also been conducted with real driving (Recarte & Nunes, 2000; Nunes & Recarte, 2002). Car drivers' pupil diameters can be measured using camera-based eye trackers (mounted either on the head or inside the vehicle). However, it is difficult to achieve reliable, useful measurements in real-world driving due to, for example, tracking problems during glances to peripheral locations (Khan & Lee, 2019) and the large impact on pupil diameter of luminance changes (Watson & Yellott, 2012).

Another fundamental mental response to cognitive demand is attention (Van Acker et al., 2018). The pupil diameter is also sensitive to changes in the level of attention, probably due to the close relation between attention and cortical arousal (Petersen & Posner, 2012), and also because the pupil has a close correlation to activity in the basal forebrain-acetylcholine system, which is associated with attention and other cognitive functions (Joshi & Gold, 2020).

In driving, not only is the level of attention important, but also where the attention is directed. Eyeblinks can be used to differentiate between visual and non-visual attention (Recarte et al., 2008). During increased visual attention, the eyeblink rate decreases (Cardona et al., 2011; Faure et al., 2016), which is assumed to occur in order to reduce the risk of missing visual information since the eyes are closed less (Recarte et al., 2008). In contrast, during increased non-visual attention (such as increased cognitive demand), eyeblink rate typically increases (Niezgoda et al., 2015; Magliacano et al., 2020), possibly to provide the brain with brief pauses from external stimuli to enhance internal processing (Nakano et al., 2013). Eyeblinks can be measured using either camera-based eye trackers or electrooculography (EOG), a technique that assesses eye movements and eyeblinks based on differences in electrical potentials in the eyes, recorded with pairs of electrodes mounted around the eyes (Schmidt et al., 2018).

In addition, more peripheral physiological functions, such as cardiovascular activity, breathing, and perspiration, are also affected by cognitive demand—although results from different studies are sometimes contradictory (for a review, see Tao et al., 2019). These functions serve to maintain the body's homeostasis and facilitate cognitive functions and behavioral responses to deal with the challenges, physical as well as mental, facing the human (Benarroch, 2011). The functions are in large part regulated by the autonomic nervous system, consisting of sympathetic and parasympathetic branches. In general, sympathetic activity facilitates emergency reactions ("fight or flight") by, for example, increasing the heart rate, directing more blood to the brain and skeletal muscles, increasing the breathing rate, and increasing perspiration in palms and feet (assumed to improve grasping; Dawson et al., 2016) (Tortora & Derrickson, 2007). Parasympathetic activity promotes "rest and digest" activities by lowering the heart rate, directing more blood to the intestines, and stimulating digestion (Tortora & Derrickson, 2007). The regulation of the autonomic nervous system is complex: the two branches can be activated separately or simultaneously to either reinforce or counteract each other (Billman, 2013), and different parts can be activated separately (Benarroch, 2011)—which means that there is great flexibility for dealing with challenges (Shaffer & Ginsberg, 2017).

The physiological responses during cognitive demand can thus have several causes. It has been suggested that the brain's increased energy and oxygen consumption in the task-

activated areas requires an increased heart rate (Fairclough & Mulder, 2012). This could explain why most studies on cognitive load report an increase in heart rate during increased cognitive demand (Hughes et al., 2019). However, the changes in neuronal activity during cognitive tasks compared to a resting state are small, and the increased demands can many times be met by redistributing resources within the brain (Shulman et al., 2004). Indeed, there are also studies that do not see an increased heart rate during increased cognitive demand (e.g., Dussault et al., 2005; Foy & Chapman, 2018).

In many situations, physiological arousal due to emotional stress has a greater impact on peripheral physiological functions than the increased energy and oxygen consumption does (Conway et al., 2013). Emotional stress can occur, for example, in situations when one experiences loss of control or failure at a task, or when test participants are being evaluated by others (Conway et al., 2013). The latter is of particular importance here, as this situation can be assumed to frequently occur in demanding experimental studies. Stress increases sympathetic nervous system activity, which is reflected in the peripheral physiological measures (Dampney, 2019).

Most peripheral physiological signals can be recorded using a few different measurement techniques. Only some of the measures and recording techniques most commonly used in experimental studies will be brought up here. Heart activity (heart rate and heart rate variability) is most commonly studied using electrocardiography (ECG), recordings of the heart's electrical activity from electrodes mounted on the upper body (Lohani et al., 2019). Perspiration is typically assessed via skin conductance, measured with electrodes mounted on the fingers, palms, or soles of the feet (Dawson et al., 2016). Breathing rate can be measured using respiratory inductance plethysmography (RIP), a technique that records chest expansions using an elastic strap around the chest (Grassmann et al., 2016). Work is ongoing to develop or refine unobtrusive recording techniques for these measures for use in driving research and vehicle applications (for reviews, see Arakawa, 2021; Leonhardt et al., 2018; Sidikova et al., 2020; J. Wang et al., 2020).

3.5.2 Other influencing factors on physiological recordings

As described above, several physiological measures respond to changes in cognitive demand and may thus be considered measures of cognitive load. However, as pointed out by Richter and Slade (2017), their use is limited by the fact that they respond not only to cognitive load but also to other mental and physical states. That is, the fact that there are multiple relationships between physiological measures and mental (and physical) states has to be acknowledged when making inferences from physiological measures (Richter & Slade, 2017).

One reason for these multiple relationships is that different mental states may rely on the same underlying mechanisms. For example, given the theory that frontal theta power reflects a need for cognitive control, the fact that increased theta power is found not only during cognitive tasks but also during sleepiness can be explained by an increased need for cognitive control to compensate for performance degradations due to sleepiness (Clayton et al., 2015). Another example is that arousal either increases or decreases in most mental states, including attention, stress, fatigue, and most emotions (Lohani et al., 2019; Pace-Schott et al., 2019). Measures that are sensitive to arousal may thus respond to changes in any of these mental

states (independent of cognitive load). Understanding and interpreting mental states in terms of their constituents or underlying mechanisms can thus reduce the number of relationships—and reduce the risk of misinterpreting responses.

Another reason for the multiple relationships is that physiological measures reflect many different functions, both related and unrelated to cognitive load. Physiological responses enable behaviors, mental and bodily functions, and adaption to environment demands. For example, breathing enables speaking, the heart rate increase during emotions that typically trigger actions (likely to prepare for physical activity; Kreibig, 2010) the pupil diameter regulates the amount of light that enters the eye for optimal vision (C. A. Wang & Munoz, 2015), perspiration helps regulate body temperature (Dawson et al., 2016), and brain activity changes with time spent awake to achieve cognitive restoration (Chong & Baldwin, 2021). Consequently, physiological responses can be expected during external (environmental), internal (bodily or mentally), and behavioral changes.

It should also be noted that some physiological responses can directly influence each other (Bari et al., 2018; Knapen et al., 2016; Quintana & Heathers, 2014), and that the size of physiological responses can depend on baseline levels (Knapen et al., 2016; Quintana & Heathers, 2014; Vetrugno et al., 2003), in addition to varying greatly between individuals (Barzegaran et al., 2017; Dawson et al., 2016; Grassmann et al., 2017; Klimesch, 1999; Mehler, Reimer, & Coughlin, 2012). For example, breathing has a direct influence on the heart rate, causing it to increase during inhalation and decrease during exhalation (Billman, 2011), and age has an influence on most physiological responses (Mehler, Reimer, & Coughlin, 2012; Umetani et al., 1998; Winn et al., 1994).

While it is important to understand the causes behind physiological responses in order to make reliable inferences about driver mental states, proper artifact handling is also a prerequisite for reliable analyses. Some artifacts come from external sources, such as cable movements or loose electrodes (Boucsein et al., 2012; Islam et al., 2016), or, in the case of camera-based eye-tracking, from lost tracking during some head, eye, or eyelid movements (Hollander & Huette, 2022). Other artifacts arise from internal, physiological sources, such as spontaneous heart rate arrhythmias (e.g., the heart skips a beat), which can cause inaccurate heart rate and heart rate variability values (Mulder, 1992). As another example, eye movements and eye blinks cause large artifacts in frontal EEG signals (Tatum, 2014). These artifacts have greatest power in low frequencies (the delta and theta frequency range) , but increase power also in the alpha and beta frequency range (Hagemann & Naumann, 2001). Changes in gaze behavior or eye blinks can thus be misinterpreted as changes in frontal EEG activity if the artifacts are not properly handled.

In summary, several physiological measures are, for various reasons, sensitive to variations in cognitive demand. The same measures are, however, also sensitive to many other things, including several other mental states as well as other internal and external factors. This complexity must be recognized when physiological measures are implemented and interpreted.

4 Methods

To accomplish the aims of this thesis, several experimental studies were conducted. The sections below will present and explain three key aspects of the studies' designs: the choice of test environment, the choice of cognitive task, and the choice of a within- or between-subjects design.

The design choices represent an effort to strike a balance between experimental control (isolating phenomena and effects of interest so they can be studied) on one hand and ecological validity (the extent to which test performance in an experiment generalizes to behaviors in real-world settings; Kihlstrom, 2021) on the other. While experimental control is important, ecological validity is crucial for ensuring that study findings are relevant to real-life applications. Any study design represents a compromise between these two. An overview of the experimental designs employed in the four papers is presented in Table 2. The way that cognitive load is assessed (or not assessed) in each paper is largely related to the choice of cognitive task and described in Section 4.2.

Table 2. Overview of research approach

	Test environment	**Type of cognitive task**	**Within/between-subjects design**
Paper I	Simulator	n-back	Between
Paper II	Simulator	n-back	Within
Paper III	Real road	Mental rotation	Within
Paper IV	Simulator	n-back	Within

Paper III differs from the other papers both in test environment and type of cognitive task because research question III (addressed in this paper) poses other requirements, as will be described further in the following sections.

4.1 Test environment

4.1.1 Laboratory

The highest level of experimental control (and typically also the lowest level of ecological validity) is obtained in lab experiments, mostly because a lab offers the opportunity to largely isolate the application of independent variables and minimize impact from confounding factors. The drawback of this approach is, of course, that the results may not be generalizable to more realistic and complex environments where the number of potential confounders is larger (Shinar, 2017).

4.1.2 Driving simulator

A driving simulator represents an intermediate step between a highly controlled lab setup and driving a real vehicle. The ecological validity can be improved since the subject is performing an actual driving task. However, the introduction of a driving task also means that the number of influencing factors increases, since the environment becomes more complex. Still, the level of experimental control is high, since the traffic scenarios to be studied can be designed with high precision and repeatability. Driving simulators thus are a great tool for studying driver behaviors or physiological responses in well-defined situations, where the goal is to explore or understand specific mechanisms (as in Papers I, II, and IV). Further, driving simulator studies make it possible to study behaviors and reactions in situations that are nearly impossible to achieve in real driving (such as the crosswind scenario in Paper IV), or that involve a safety risk (such as the critical event in Paper I) (Shinar, 2017). For Papers I, II, and IV, the choice to collect data in driving simulators was made to achieve a high level of experimental control over the traffic scenarios to be studied.

However, one must acknowledge that driving a simulator is not the same as driving a real vehicle; therefore, just as for in-lab experiments, it is not certain that effects seen in simulators can be generalized to real life (de Winter et al., 2012). In a review of validation studies comparing different measures (primarily behavioral) taken during simulated and real road driving, Wynne et al. (2019) concluded that there were large differences in simulator validity between studies. The differences depended on the simulators tested (it is noteworthy that simulators with higher fidelity, such as a large field of view and a moving base, did not always show higher validity than lower-fidelity simulators) as well as on the dependent variables used in the validation. Subjective ratings (Diels et al., 2011; Fors et al., 2013; Lobjois et al., 2021), response task performance (Lobjois et al., 2021), and eye blink frequency (Lobjois et al., 2021) suggest that driving a simulator is more cognitively demanding than driving a real car. HR levels have, in turn, been found to be lower in simulated compared to real driving (Johnson et al., 2011; Reimer & Mehler, 2011), a result though to reflect lower levels of perceived risk in the simulators. Other studies have, however, not found a significant difference in HR between simulated and real driving (Milleville-Pennel & Charron, 2015; Mueller, 2015).

As discussed in Section 3.2.1, the cognitive control hypothesis (Engström et al., 2017) emphasizes the need for realistic driving tasks when behavioral effects of cognitive load are to be studied (as in Papers I and II). Since the hypothesis suggests that the effects of cognitive load depend on whether the (driving) behavior in question requires cognitive control, it is important that the driving tasks resemble real-life driving and require the same amount of control. Therefore, a great deal of effort was put into making the driving tasks in Papers I and II as realistic as possible. An advanced moving-base simulator was used for the experiments, and the driving environment largely mimicked a real road in Sweden. Importantly, the experimental drives were conducted as continuous drives of approximately 40 minutes, during which the drivers were not interrupted except when asked to perform cognitive tasks. Also, self-ratings and secondary task performance measures were not used to assess cognitive load, since they can alter or interrupt the driving task (as described in Section 3.5).

4.1.3 Test track

Having participants drive real vehicles rather than simulators improves the ecological validity of a study further. Moreover, experiments conducted on closed test tracks rather than in real traffic can still achieve a high level of experimental control. It is, however, expensive and difficult to create scenarios that involve other road users or complex traffic environments which are both realistic and safe. To guarantee safety, balloon vehicles and pedestrian or bicyclist dolls are typically used to represent other road users (e.g., Boda et al., 2018; Victor et al., 2018). Unfortunately, the artificiality of these elements may influence driver behaviors and responses, limiting the generalizability of the results.

4.1.4 Real traffic

Moving from the test track to real traffic increases the ecological validity further, but with a considerable loss of experimental control. For example, participants cannot be exposed to safety-critical events for safety reasons, and the control over the behavior of other road users is extremely limited, making it difficult to study behaviors and responses in specific scenarios. Yet, for the research question posed in Paper III to be answered accurately, it was necessary that the participants' perceived risk levels corresponded to those in real life. It was therefore decided to conduct the experiment in real traffic. It is important to bear in mind that, although the environmental context is fully natural, test participants in studies on real roads are still often being explicitly observed, so the situation may not actually be very naturalistic. In Paper III, for example, there were test leaders in the vehicle, the participants wore test equipment, and they were given a specific task to perform. Thus, participants' behavior and responses even in studies on real roads may still differ from those in everyday driving.

Full ecological validity can probably only be achieved in observational or naturalistic driving studies which collect data unobtrusively with vehicles that are being used in everyday driving in real traffic. Since the experimental control in these types of studies is close to zero, very large amounts of data must be collected to achieve a sufficient sample of, for example, safety-relevant events for statistical analyses (Dingus et al., 2016). Further, specific situations will rarely repeat (since they are driven by chance, not design), and knowledge about drivers' mental state in any particular situation is usually lacking. These limitations make naturalistic driving data unsuitable for the aims of this thesis.

4.2 Cognitive tasks

Understanding how cognitive tasks affect driver behaviors and traffic safety requires utilizing different types of tasks for different purposes, just as different test environments need to be used. In most cognitive tasks that drivers engage in during everyday driving, such as cell phone conversations and controlling the infotainment system using voice control, the variation in cognitive load can be large. This variation is partly explained by the large variety in task content—and also by the fact that, in most tasks, drivers can adapt their engagement and effort, and thereby their cognitive load, to fit the situation as a whole (Lee, 2014). Drivers may also adapt how (Tivesten & Dozza, 2014) and when (Tivesten & Dozza, 2015) to engage in a specific task, and employ risk-compensating strategies, such as lowering their speed and increasing their time headway during task execution (Wijayaratna et al., 2019). To

understand the effects that cognitive tasks have on traffic safety it is thus necessary to allow for drivers' adaptation to the situation as a whole, and high ecological validity is consequently required in both the task and the context.

In reductionistic studies that investigate specific effects of cognitive task engagement (rather than task engagement as a whole), there may be a need to control the level of cognitive load more closely, making naturalistic tasks unsuitable. This is also the case in studies that search for reliable measures of cognitive load (measures which, once determined, can enable less controlled studies). For these reasons, artificial tasks with a high degree of experimental control were the choice in all studies in this thesis work.

In Paper III, task engagement had to continuously generate not only cognitive load, but also visual and motoric load, and task performance had to be measured in order to study the participants' ability to disengage from traffic and perform a demanding task while traveling in AD mode. Thus, a task on a handheld tablet was created that had participants perform mental rotations of complex geometric shapes. That the task was indeed cognitively loading was confirmed with self-reports. Although this was not a task that drivers regularly encounter in real life, it enabled the research question to be answered.

For the studies in Papers I, II and IV, requirements on demand continuity and level of cognitive load were much higher than in Paper III's study, since the specific behavioral and physiological effects of cognitive load were studied with high temporal resolution. Also, self-reports and secondary task performance measures could not be used, as they might have altered or interrupted the driving task (as discussed above). Thus, to ensure that a continuous cognitive load was applied to the drivers and that the level of load could be systematically varied, a task frequently used in research studies with a well-established effect on cognitive load (Jaeggi et al., 2010) was employed: the n-back task. Employing this task thus made it reasonable to presume the existence of cognitive load during task performance without having to actually measure it.

In the n-back task, the participants are presented with a stream of stimuli at a fixed pace and asked to determine whether each stimulus matches the one presented *n* stimuli ago (Mehler et al., 2011). The level of difficulty is increased by increasing *n*. Because the participants are asked to answer as fast as they can (in our case by pressing a button when a match was detected), the task is performance maximizing, meaning there is less room for variation in task engagement and effort compared to more naturalistic tasks. The n-back task thus allows for high experimental control, which in turn enables detailed analyses of behavioral mechanisms and physiological responses.

4.3 Within- or between-subjects design

Another important choice in experimental studies is between a within- or between-subjects design. In a within-subjects design, the participants are exposed to all the different experimental conditions, such as the different levels of cognitive load (Shinar, 2017). In a between-subjects design, different groups of participants are instead exposed to different experimental conditions, so that each participant is only exposed to a single condition, such as a single level of cognitive load (Shinar, 2017).

Within-subject designs require fewer test participants and the statistical power is often larger, since individual differences have less impact on the results (Keren, 2014). However, a major concern in within-subjects designs is the risk of order effects (Shinar, 2017). In driving studies looking at the effects of cognitive tasks, participants may, for example, exhibit a learning effect over time both in the cognitive task (Wheatley et al., 2018) and in the driving task (Ronen & Yair, 2013; Rosey & Auberlet, 2014), so that the tasks become less challenging over time. Other examples of order effects include increased fatigue and/or less stress over time (Belyusar et al., 2015). As a result of these effects, the cognitive load may thus both decrease and change in composition during the course of the experiment. While many order effects can be reduced or removed by practice (Haith & Krakauer, 2018), all order effects cannot be fully removed (Keren, 2014). Instead, the order of experimental conditions is typically balanced across participants so that the order effects are (hopefully) equally spread across conditions (Keren, 2014).

Another related issue with within-subjects designs is that, based on what has happened previously, participants may form hypotheses about what will happen next and adapt their behaviors accordingly (Keren, 2014). Consequently, if scenarios are repeated, participants are likely to anticipate events. Recall that the cognitive control hypothesis (Engström et al., 2017, see Section 3.2.1) suggests that the effect of cognitive load will differ in an unexpected event in which some stimulus (e.g., visual looming) automatically triggers a response compared to in an event where a response is either initiated or enhanced by cognitive control (such as when an event is anticipated due to repetition). In Paper I, the research question required that participants responded to an *unexpected* lead vehicle braking event. Therefore, this event could not be repeated and thus a between-subjects design was necessary.

In Papers III and IV, the effects of task repetition were explicitly tested, making a within-subjects design the natural choice. A within-subjects design was also employed in Paper II, in that case to improve the statistical power. The task and scenario repetitions may have led to learning effects with concomitant changes in drivers' gaze behavior, but this possibility was not explored.

5 Summary of papers

Paper I – Effects of Cognitive Load on Response Time in an Unexpected Lead Vehicle Braking Scenario and the Detection Response Task (DRT)

Introduction. Many experimental studies have demonstrated increased response times to various stimuli during cognitive activities. This has led to concerns that cognitive activities will also increase response times in critical situations in real-life driving. A commonly used response time task which consistently shows increased response times during cognitive activities is the ISO standardized Detection Response Task (DRT). However, a few studies have been performed in which the driver's responses were triggered solely by visual looming, and no effect of cognitive activities on response times was found.

Aim. The aim of this paper was to see if the same cognitive task had similar effects on response times in the DRT as in an unexpected lead vehicle braking scenario with strong visual looming.

Method. The study consisted of two experiments. In Experiment 1, 16 participants drove a fixed-base driving simulator on a four-lane highway at approximately 90 km/h. In addition to simply driving (baseline), they performed an audio version of the cognitively loading n-back task at two levels of difficulty, 1-back and 2-back. The tasks were performed with and without concurrent execution of the tactile DRT.

In Experiment 2, 24 participants drove a moving-base driving simulator on a two-lane rural road at approximately 80 km/h. After approximately 40 minutes of driving, the participant's vehicle was overtaken by another car, which suddenly braked in front of the participant. The scenario was designed so that visual looming appeared as soon as the overtaking car started braking. The participants were either just driving, or involved in the 1-back or 2-back task when the overtaking car started braking.

Results. In Experiment 1, DRT response times increased significantly during cognitive task engagement; being on average 0.33 s in the baseline condition, 0.47 s in the 1-back condition, and 0.50 s in the 2-back condition, $F(2,45)=16.24$, $p<.0001$. In Experiment 2, there was no effect of the same cognitive tasks on brake response times; being on average 1.00 s in the baseline condition, 0.97 s in the 1-back condition, and 0.98 s in the 2-back condition, differences which were not statistically significant.

Discussion. The same cognitive tasks had different effects on the response times in the DRT and in the unexpected lead vehicle braking scenario. This finding can potentially be explained by the cognitive control hypothesis (Engström et al., 2017), which says that cognitive load selectively impairs driving subtasks that rely on cognitive control but leaves automatic performance unaffected. According to the hypothesis, since the DRT was a relatively novel task for the participants, it required cognitive control; hence its performance was negatively affected by cognitive tasks. Responses to visual looming are, however, innate and automatic and thus, according to the hypothesis, should not be affected by cognitive tasks. The results suggest that it is inappropriate to generalize the effects of cognitive tasks seen in cognitively controlled response tasks (such as the DRT) to critical events where drivers can respond to visual looming.

Paper II – 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 cognitive task and the traffic environment

Introduction. Visual information from the traffic environment is necessary for safe driving. Numerous studies have found a gaze concentration effect with more on-road glances during cognitive task execution. It is, however, not clear why this effect occurs, or how the timing of glances towards relevant information in off-path locations is affected by cognitive tasks.

Aim. The aim of this paper was to better understand the gaze concentration effect by studying the relation between gaze direction, cognitive task engagement, and the traffic environment on a high-resolution time scale in two traffic scenarios with potential threats in off-path locations.

Method. Thirty-six participants drove on a rural road in a moving-base driving simulator. The two scenarios, a four-way intersection scenario and a hidden exit scenario, were repeated four times each during the 40-minute drive. The scenarios were designed to contain potential threats in off-path locations. When encountering the scenarios, the participants were either engaged in a cognitive task (the 1-back or 2-back task) or they were just driving. The drivers' glance behavior was recorded using an eye-tracker.

Results. The drivers adapted their visual behavior to the traffic scenarios by decreasing the number of on-path fixations and directing their gaze towards relevant information in off-path locations. When the drivers were cognitively loaded, the percentage of the total time that they looked on-path increased from 49.7% (intersection) and 64.7% (hidden exit) in the baseline condition to 56.6% and 72.1% in the 1-back condition and 66.4% and 76.5% in the 2-back condition. More detailed analyses revealed that gaze shifts from an on-path to an off-path location were inhibited by cognitive load, while gaze shifts in the other direction remained unaffected. Consequently, on-path glances increased in duration, while off-path glances decreased in number. The timing of the remaining glances towards the off-path locations was, however, unaffected by cognitive load.

Discussion. The observed gaze concentration effect could be explained by the selective inhibition of gaze shifts from on-path to off-path locations. Assuming that looking on-path is an automatized behavior (since it is, by far, car drivers' most frequent gaze direction), this finding can potentially be explained by the cognitive control hypothesis (Engström et al., 2017). Furthermore, given that cognitive load seems to have inhibited rather than delayed gaze shifts (unlike response times in many response tasks), and that gaze direction control itself seems unaffected by cognitive load, it appears that cognitive load has different effects on different mechanisms.

Paper III – Drivers' ability to engage in a non-driving related task while in automated driving mode in real traffic

Introduction. With the introduction of systems for AD, drivers will be free to disengage from driving and engage in non-driving-related tasks (NDRTs), such as working, while on the move. Having this possibility is considered one of the main advantages with AD. However, while the interest in engaging in NDRTs while traveling is confirmed, there is a lack of empirical studies investigating to what extent people are in fact able to engage, and maintain engagement, in NDRTs while traveling in AD mode in real traffic.

Aim. The aim of this paper was to investigate how well drivers are able to engage in an NDRT while traveling in AD mode in real traffic.

Method. Ten participants traveled in AD mode in real traffic via a Wizard of Oz platform. They were informed that Volvo Cars were responsible for the safety once they had activated the AD function. While traveling in AD mode, the participants performed an NDRT which was designed to be visually and cognitively demanding and to require manual interaction: it consisted of performing mental rotations of polygons on a tablet. The drivers' task engagement was assessed by their shifting of gaze direction from the road ahead to the NDRT, by their performance in the NDRT (number of correct answers and response times) while traveling compared to while in an office environment, and from their self-reported experience (from questionnaires and interviews) of engaging in the NDRT.

Results. The results show that when the drivers performed the NDRT while in AD mode, their gaze direction shifted from the road ahead towards the NDRT to a great extent; gaze was directed towards the road ahead for less than 20% of the time in all drives, and for less than 2% of the time in 50% of the drives. There were no significant differences in their performance in the NDRT when in the car compared to in the office, nor in their rated experiences (although there was a tendency for perceived task demand to be rated higher and engagement lower in the car). Yet, many participants reported in the interviews that they noted and reacted to changes in the environment and to sudden vehicle motions. Some participants were surprised by their own ability to easily disconnect from driving.

Discussion. This study extends previous research by demonstrating that drivers are able to engage in an NDRT to a great extent while in AD mode in real traffic. This finding supports the promise that future cars with systems for AD will be able to "free up time" that drivers can use to engage in non-driving-related activities while traveling.

Paper IV – Let complexity bring clarity: A multidimensional assessment of cognitive load using physiological measures

Introduction. Despite plenty of studies on the topic, the effects of secondary task induced cognitive load on driver behavior and traffic safety are unclear and in need of further investigation. For this purpose, reliable measures of cognitive load are needed. Physiological measures appear useful for this purpose since they can provide continuous recordings of drivers' mental state without interfering with the driving task. There are, however, a few issues related to their use in this context which are typically overlooked and which may cause inaccurate inferences and generalization. First, cognitive load consists of multiple mental responses (cognitive load components) to cognitive task demands and can therefore take many different forms—yet, researchers treat it as if it is unidimensional. Second, task-induced cognitive load does not occur in isolation; rather, it is part of a complex mental response to the situation as a whole. Third, physiological measures are affected by multiple mental and physical states, limiting the inferences that can be made from them individually.

Aim. The aim was to demonstrate how the measurability of cognitive load can be greatly improved if the above-mentioned issues are acknowledged. This aim was accomplished by studying multiple mental responses using multiple physiological measures and independent variables.

Method. 70 participants drove on a rural road in a moving-base driving simulator for approximately 40 minutes. Three different traffic scenarios were repeated four times each. When encountering the scenarios, the participants were either engaged in a cognitive task (the 1-back or 2-back task) or they were just driving. The 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 taken and analyzed in relation to cognitive task demand, scenario repetition, and type of traffic scenario. Cognitive load components and other coinciding mental responses were assessed by considering the response patterns of multiple physiological measures.

Results. The cognitive tasks had a significant effect on heart rate, heart rate variability, breathing rate, skin conductance, pupil diameter, and eye blink rate, but not on eye blink duration, EEG alpha power, or EEG theta power. Several of the measures were also sensitive to scenario repetition and/or type of traffic scenario, but, importantly, the effects were not the same for all measures.

Discussion. Several comparisons have been made between uni- and multi-dimensional interpretations of physiological responses. This way, it was demonstrated how the construct validity of cognitive load was improved by acknowledging the fact that multiple mental responses occurred during the course of the experiment. Also, the concurrent analysis of multiple measures made the measurements more diagnostic—that is, better able to distinguish between different mental responses. This improvement increases the measurements' overall external validity, as the risk of making incorrect inferences is reduced. As more diagnostic and valid measures of cognitive load components become available, the effects of cognitive task demands on traffic safety can be better understood, and possibly mitigated.

6 Discussion

This chapter addresses the main findings from the papers in this thesis with respect to the thesis aim and research questions (stated in Chapter 2). The first three sections will discuss the observed effects of cognitive tasks on driver behaviors and the implications of these effects for traffic safety in safety-critical events (Section 6.1), in routine driving (Section 6.2), and in the pre-trip phase when decisions about how to travel are being made (Section 6.3). Section 6.4 will discuss why changing the way we look at cognitive load from a unidimensional to a multidimensional view (acknowledging the fact that multiple mental responses occur during task engagement) is of key importance for making physiological measures truly useful in mental state assessments. Section 6.5 will then discuss what acknowledging multiple mental responses to task engagement means for studies on driver behavior and traffic safety. Section 6.6 will discuss different areas of application for physiological measures in traffic safety. Lastly, suggestions for future work in this area will be discussed in Section 6.7.

The intent of this work is to contribute to a more detailed understanding of the complicated relationships between cognitive task involvement and driver's mental responses, behaviors, and physiological responses. The desired result is contribution to a nuanced discussion, rather than absolute answers in terms of specific safety consequences or measurement guidelines.

6.1 Response times in critical events

The results from Paper I (Research Question 1: How do cognitive tasks affect drivers' response times in an unexpected safety-critical scenario?) suggest that cognitive tasks will not slow down driver responses in critical scenarios, as long as automated responses are sufficient to resolve the situation. Only in situations where cognitive control is required for an appropriate response are cognitive tasks expected to increase risk. Specifically, it was found that the same cognitive task had different effects on response times depending on the stimulus to which the driver was to respond. Response times in the DRT were increased during cognitive demand, in line with previous findings (e.g., Bruyas & Dumont, 2013; Conti et al., 2012; Engström, Larsson, et al., 2013), while response times in a critical and unexpected lead vehicle braking scenario were unaffected by the same cognitive tasks. Importantly, the lead vehicle braking scenario was designed so that visible looming of the lead vehicle appeared as soon as it started braking, so the drivers presumably responded automatically to the looming cues. According to the cognitive control hypothesis (Engström et al., 2017), this is why cognitive load did not have an effect on response times.

These results clearly demonstrate that the effects of cognitive tasks seen in cognitively controlled response tasks, such as the DRT, should not be generalized to critical situations in which drivers can respond to visual looming, as has been done previously (e.g., Strayer & Fisher, 2016; Merat & Jamson, 2008). Since the results support the cognitive control hypothesis, it seems instead that cognitive tasks should only increase the risk of accidents in situations where drivers do not have a suitable automatized behavior to fall back on (Engström et al., 2017). The effect on traffic safety hence depends both on the driver's

experience (what behaviors s/he has automatized) and the specific situation. In situations where cognitive control is involved, response times are likely to increase, just as the DRT response times did. Such a situation could, for example, involve negotiating with multiple road users in a complex intersection, where fast decisions are necessary. Behavioral adaptation can also be expected to suffer in any situation where cognitive control is required for an appropriate, although not necessarily fast, response, such as when a potential conflict can be predicted using cues in the traffic environment (see Baumann et al., 2008 and Muttart et al., 2007). This expectation is discussed further in Section 6.2 below.

It should be noted that while there are several studies that consistently report that cognitive load causes increased response times in tasks which rely on cognitive control (such as the DRT), there are very few studies which have explored the effects on responses to visual looming alone. Further studies are thus needed to confirm the results in Paper I.

6.2 Glance behaviors in routine driving

It was found in Paper II (Research Question 2: How do cognitive tasks affect drivers' glance behavior in routine driving?) that gaze concentration, which typically occurs during engagement in cognitive tasks (e.g., Recarte & Nunes, 2000, 2003; Reimer et al., 2013; Victor et al., 2005; Y. Wang et al., 2014), seems to stem from gaze-shift inhibitions, rather than gaze-shift delays. The drivers looked more towards the future path (i.e., they displayed gaze concentration) during cognitive task execution compared to when they were just driving (baseline) in two traffic scenarios which both had relevant information in off-path locations: a four-way intersection scenario and a hidden exit scenario. This finding is in line with the results from many previous studies, most of which describe the effects of cognitive tasks on drivers' visual behavior on highways and main roads (Recarte & Nunes, 2000, 2003; Reimer et al., 2013, 2012; Reimer, 2009; Victor et al., 2005; Y. Wang et al., 2014), although some were conducted in more complex traffic environments (Harbluk et al., 2007; Biondi et al., 2015).

Interestingly, when glance behaviors were studied at a high temporal resolution in Paper II, it was found that the timing of glances towards relevant areas in off-path locations was unaffected by cognitive demand. In other words, the gaze patterns in relation to the traffic environment were unaffected by cognitive demand in terms of timing, but changed so that the proportion of on-path fixations was larger. Note that cognitive demand did not cause a delay in glances towards off-path locations (as speculated by Lehtonen et al., 2012), a result which would have been consistent with the well-known effect of delayed response times in tasks requiring cognitive control. Instead, it appears that cognitive load leads to some off-path glances being omitted. Moreover, the risk of such "absent glances" did not seem to depend on the traffic environment.

Further, it was found in Paper II that the effect of cognitive demand on visual behavior was dependent on gaze direction. While gaze shifts from an on-path to an off-path direction were inhibited by cognitive demand, gaze shifts in the opposite direction (off-path to on-path) remained unaffected. This resulted in fewer off-path glances and longer on-path glances: that is, gaze concentration.

One way to understand this finding is, again, in terms of the cognitive control hypothesis (Engström et al., 2017). If we assume that drivers automatically direct their gaze towards the future path (because that is the most frequent and consistent gaze direction needed for the most frequent and consistent task, lane keeping), then eyes-on-path becomes the automated and default gaze location. It would follow that moving the gaze to an off-path location requires cognitive control (as has been speculated by Lehtonen et al., 2012). That is, drivers who have an on-path gaze direction during increased cognitive load are less likely to switch to an off-path gaze direction, since that would require cognitive control. Drivers who, on the other hand, have an off-path gaze direction during increased cognitive load can switch back to the default on-path gaze direction automatically, and cognitive load would not disrupt that gaze switch. Thus, gaze concentration can be viewed as an automatic behavior in which gaze reverts to the default location. This behavior is likely automatized over time, through successful employment of the neural pathways (Cohen et al., 1990) used during the extensively practiced lane keeping task.

The effect of gaze concentration on traffic safety is most likely situation-dependent. The gaze concentration effect in drivers was originally observed in studies conducted on highways and larger roads (e.g., Recarte & Nunes, 2000, 2003; Reimer et al., 2013, 2012; Reimer, 2009; Victor et al., 2005; Y. Wang et al., 2014). On such roads, there is not much safety-relevant information in off-path locations and hence, spending more time gazing at the road ahead is likely to be safety-beneficial. In fact, as mentioned in Section 3.3, the gaze concentration effect has been suggested as an explanation for the decreased risk of rear-end collisions during cell phone conversations (Dingus et al., 2019; Victor et al., 2015). However, in more complex scenarios such as intersections, it is unlikely that gaze concentration is safety-beneficial. Instead, a decreased proportion of eye fixations towards (potentially) safety-relevant objects in off-path locations is likely to reduce the driver's situation awareness, making him/her less prepared to act if a potential threat turns into an actual threat (Harbluk et al., 2007; Biondi et al., 2015). It is thus possible that if another vehicle had entered the main road in the hidden exit scenario in Paper II, the drivers would have noticed it later and/or reacted to it later. This deserves investigation in future research.

The results in Paper II call for further research to understand the interaction between cognitive tasks and the task of driving. Because the drivers did not appear to compensate for "absent glances" to off-path locations by performing them later, the timing between variations in cognitive load and intended gaze shifts is significant from a traffic safety perspective. The probability that a driver will fail to scan an important area might thus depend on the timing between intended gaze shifts to that area and peaks in cognitive demand. Just as the timing between off-path glances and lead vehicle braking events are decisive for the risk of rear-end collisions (Victor et al., 2015), so could the timing between cognitive task engagement and intended gaze shifts be decisive for the risk of missing relevant information in off-path locations.

For the design of driver support systems, it is very important to know when system interventions should occur. A system which sends warnings too late is, of course, less effective. But a system which sends warnings too early (for example, a forward collision warning system that warns an attentive driver of an impending collision a second before the driver was already going to brake) is likely to be perceived as annoying and turned off by the driver. However, depending on driver state, drivers might need support at different times.

Imagine a system that could help drivers achieve situation-appropriate scanning behavior in intersections. First of all, such a system would need to take the traffic environment, including other road users, into account when deciding what is appropriate glance behavior for the driver. Imagine that the system detected that the driver did not look towards a crossing cycle path as expected and needed, and that the system also knew that the driver was engaged in some cognitive task. If cognitive load causes off-path glances to be omitted rather than delayed, the system should not wait to remind the driver to scan the environment. If, on the other hand, glances are merely delayed, there is reason for the system to wait with the intervention to avoid annoying the driver. Therefore, in order to improve the systems' capacity to support safe driving, more research is needed to confirm whether the findings in Paper II persist in other studies and scenarios.

6.3 Pre-trip decisions to use autonomous driving systems

AD seems to be able to free up time for working or engaging in other NDRTs while traveling, just as people want (Hecht et al., 2020; Pfleging et al., 2016; Winkeler et al., 2019). Having the possibility to engage in NDRTs while traveling will therefore likely be a compelling argument for deciding to use AD systems, once available. The results in Paper III show that the drivers were indeed able to engage in the NDRT. They directed their gaze towards the NDRT for the vast majority of time and performed it as well in the car as they did in an office environment. The drivers also reported that they were able to concentrate and become absorbed in the task, although it should be kept in mind that there were large individual differences in how task engagement was experienced. While some participants experienced the situation as rather stressful and stated that they tried to divide their attention between the NDRT and the traffic environment, several reported being pleasantly surprised by their experience of letting go of driving and engaging in the NDRT.

Somewhat surprisingly, even the most skeptical participant was able to perform the NDRT well and (to some extent) let go of driving. This participant described herself as a late adopter who never used any systems for assisted driving and never used a cell phone while driving. She also said in the interview that she did not want to disengage from of driving; still, during the first test occasion she only looked at the road ahead 15% of the time, and her longest off-path glance was 11 s. On the second test occasion, she reported feeling more relaxed and said, "Basically I did not look at the road hardly at all." Here, the percentage of time she looked at the road ahead was 12% and the longest off-path glance was 10 s. To put these numbers in context, it has been shown that drivers who engage in NDRTs during manual and assisted driving very rarely look away from the forward roadway for more than three seconds; a naturalistic driving study by A. Morando et al. (2020) found that only 1% of off-path glances had a duration beyond 2.6 s and 2.8 s in manual and assisted driving, respectively. The skeptical participant, although she said that she tried to monitor the traffic, did indeed make use of the AD function to let go of driving and engage in the NDRT. Clearly, not only early adopters and AD enthusiasts are able to hand over control to the car.

It should be noted that in Paper III, test participants were explicitly asked to perform an artificial task while traveling in AD mode in real traffic while test leaders were present in the rear seat. This experimental design addressed Research Question 3 (To what extent are

drivers able to engage in non-driving-related tasks while traveling in AD mode in real traffic?); however, many questions remain about the manifestation of task engagement in future vehicles offering AD. For example, drivers such as the skeptical participant described above might choose not to engage in NDRTs if not explicitly asked to do so. Finding out how future systems will be used is complicated for several reasons, not the least because behaviors tend to change over time as the systems are being used—referred to as behavioral adaptation (Sullivan et al., 2016). In Paper III, the drivers participated in the study on two occasions, with approximately one week in between. This design meant that the first encounter and any early learning effects could be accounted for. These effects were rather small, but there were indications both in the interviews and the glance behaviors that drivers became more relaxed and could engage more in the NDRT on the second occasion. Earlier studies have also demonstrated that drivers became more relaxed and changed their behaviors with repeated exposure to AD (Andersson et al., 2018; Large et al., 2019). However, behavioral adaptation to AD systems most likely continues for much longer when drivers use the systems in their everyday driving than what has yet been studied. A naturalistic driving study by Dunn et al. (2019) found that drivers who used a car with systems for assisted driving for one month exhibited behavior that was different than that of drivers who already had the systems in their cars (and thus were more accustomed to them) and were observed for approximately one year. When the systems were activated, the drivers who were less experienced engaged less in NDRTs and looked away from the forward roadway less than when the systems were available but not activated; the more experienced drivers did the opposite, engaging more in NDRTs and looking away from the forward roadway more. Naturally, naturalistic testing of systems which are not yet fully developed (such as the system for AD employed in Paper III) is difficult, so this type of long-term behavioral adaptation is hard to study empirically before products are actually on the market.

Drivers' experiences when using a system determine how behavioral adaptation develops and whether the system is appreciated and will be used (Dunn et al., 2019). An interesting finding in Paper III was that some of the test participants stated that they felt that task engagement helped them let go of driving and trust the system. It therefore might be beneficial to encourage involvement in NDRTs to create trust in the system.

When drivers do other things behind the wheel (such as working or sleeping) during AD, changing their position in the car for improved comfort and task performance is a likely effect (Köhler et al., 2019). In fact, a participant in the Paper III study noted that the space in the driver seat was tight and not an optimal working environment when, for example, using a computer. Therefore, research and product development is also needed (and, indeed, underway) to understand how drivers wish to sit, so that more seating positions are available in vehicles with AD systems—without compromising occupant safety in the event of an evasive maneuver or crash (see e.g., Jorlöv et al., 2017; Leledakis et al., 2021).

6.4 Assessing cognitive load using physiological measures

The inconsistent effects of cognitive tasks on physiological measures in different studies (see Tao et al., 2019, for a review) can largely be explained by the fact that the studies assess cognitive load on a unidimensional scale, even though it is a multidimensional response (see

Section 3.1). If we stop trying to assess cognitive load as a singular entity, and instead focus on more specific and measurable mental responses and their neurological underpinnings, our chances of performing useful and reliable assessments will increase significantly.

It was demonstrated in Paper IV (Research Question 4: How can cognitive load in car drivers be assessed using physiological measures?) that cognitive task demand gave rise to different physiological responses as the mental responses changed during the course of the experiment. Consequently, by acknowledging multiple different—and often coinciding—mental responses to multiple influencing factors (including task demands and environmental factors such as traffic environment complexity), mental state measurability can be greatly improved.

It follows that our chances of assessing cognitive load in a useful and reliable way will increase if we can understand which mental responses are actually reflected by the physiological responses. In other words, it seems reasonable to stop trying to assess cognitive load per se and instead focus on more specific and measurable mental responses. For example, rather than interpreting increased heart rate as a sign of cognitive load, it may instead prove more useful for assessing emotional stress or physiological arousal (Conway et al., 2013), which may or may not occur during cognitive task engagement.

It is also important to remember that physiological measures are sensitive to multiple mental responses (Richter & Slade, 2017). To reduce the number of relationships between mental and physiological responses, concurrent analyses of multiple physiological measures are needed. Studying multi-measure response patterns rather than single-measure responses makes it possible to exploit both the similarities and differences between the different measures. For example, as demonstrated in Paper IV, visual and non-visual attention can be differentiated when pupil diameter (which increases in response to increased attention regardless of modality) and eye blink rate (which decreases during increased visual attention but increases during increased non-visual attention) are analyzed jointly (Recarte et al., 2008). The distinction between visual and non-visual attention can be of great importance when assessing a driver's mental state, since the type of attention can indicate the difference between a distracted driver and an attentive one. If gaze direction is also measured, it is possible to also determine whether visual attention is directed towards safety-relevant areas in the traffic environment (indicating an attentive driver), or towards something un-related to driving (indicating a visually distracted driver).

Studying multi-measure response patterns instead of single-measure responses also reduces the number of correlations between physiological measurements and different mental states. This means that the measurements' context dependence is reduced, since fewer factors affect the measurements. Consequently, the risk of making incorrect inferences from observed responses is reduced and the measurements' external validity is improved.

The use of multiple measures has indeed been encouraged for a long time (see, e.g., de Waard, 1996). In fact, most research seeking physiological indicators of cognitive load, especially that which employs machine learning to develop cognitive load classifiers, does indeed include multiple measures in the analyses (e.g., Chihara et al., 2020; Murphey et al., 2019; Putze et al., 2010). However, one cannot develop cognitive load classifiers simply by using many physiological measures if the fundamental problem—attempting to assess a multidimensional construct on a unidimensional scale—remains. The fact that multiple mental responses co-occur needs to be acknowledged in order for the classifiers to function in

different settings. If a classifier is trained on a limited range of cognitive load types (i.e., not all mental responses that can comprise cognitive load are included in the training data), then the classifier is likely to struggle to detect cognitive load when used in a different setting with a novel (for it) type of cognitive load. This difficulty is likely to occur if classifiers are trained in controlled experiments in which participants experience relatively high levels of stress and effort, and are then used in naturalistic settings—where cognitive tasks are self-initiated and the resulting cognitive load is lower in stress and effort because the driver is more relaxed.

It should be noted that there are indeed studies that acknowledge that different measures are sensitive to different influencing factors, and that more than one type of cognitive load occure in response to cognitive task manipulations (Patel et al., 2018). A meta-analysis by Hughes et al. (2019) of cardiac responses to perceptual and cognitive task manipulations concluded that different measures were sensitive to different influencing factors (e.g. event rate and task duration). The authors suggested that different measures should therefore be used to assess cognitive load, depending on the type of task at hand. However, even if they acknowledged that different types of cognitive load can occur, they still refered to all of them as "mental workload". This is unfortunate, because if the mental responses that are de facto being measured are specified, rather than grouped together under an umbrella term like cognitive load (or similar), our chances of finding measures that work in many settings would increase. Also, specifying the individual mental responses rather than grouping them together would greatly facilitate our exploration of the effects of cognitive tasks in different situations (including their effects on driver behaviors), and subsequently help us determine which mental responses are of greatest importance for traffic safety.

Importantly, when acknowledging the mixture of multiple, coexisting mental responses in a specific setting, it is unnecessary to decide which mental responses should be included in the cognitive load construct and which should be excluded (which, as mentioned in Section 3.1, is difficult). For example, in Paper IV it appeared that in the early task repetitions the participants were experiencing high emotional stress (indicated by increased heart rate) and high cortical arousal (indicated by increased pupil diameter), while in the later repetitions they were more relaxed, but still increased their cortical arousal in response to the task demand. Depending on how cognitive load is defined, and which measures are used to assess it, these results could be interpreted differently: either the level of cognitive load was constant throughout the experiment (if only the cortical arousal reflected in the pupil diameter is considered), or the level of cognitive load decreased during the course of the experiment (if stress reflected by heart rate is considered). It seems meaningless to choose between these interpretations when both measures can be informative if multiple coexisting mental responses to cognitive task demand are assessed individually. That is, during the course of the experiment, drivers became more comfortable and less stressed when doing the tasks, but they kept investing similar amounts of cognitive effort in the task performance.

6.5 Considering multiple mental responses when studying driver behaviors

Importantly, the different mental responses (e.g., stress and cognitive effort) that occur when drivers engage in cognitive tasks can have different effects on driver behaviors. Different studies of cognitive load are thus not necessarily comparable just because they include some

sort of cognitive demand, and a large measure of caution must be applied when comparing and generalizing results.

The way the effects of cognitive tasks on driver behaviors has been studied and interpreted in Papers I and II is, however, in line with most research today, as well as with the cognitive control hypothesis. That is, when cognitive task demand was manipulated, cognitive load was assumed to increase or decrease on a unidimensional scale, and possibly induce changes in driver behaviors. This approach cannot determine whether the observed effects would be the same if the mental responses to the cognitive demand were different.

According to Engström et al. (2017), all cognitive tasks should, however, have the same effect on driver behaviors (in terms of drivers falling back into automatized behaviors) since the effect is attributed to their common denominator: the prefrontal cortex activity which comprises cognitive control. It should thus not matter which other mental responses co-occur with this prefrontal cortex activity. Based on the results in Papers I and II, it is not possible to confirm or reject this assumption. That is, we currently do not know what would have happened in the lead vehicle braking scenario in Paper I if not only the drivers' level of cognitive control but also their level of, for example, stress had been high, since the way the scenario was implemented the drivers' level of stress was de facto low (as confirmed in Paper IV).

There should, however, be little doubt that other mental responses which can also result from cognitive task demands (along with cognitive control) can also affect driver behaviors. Some mental responses, such as negative emotions (Oei et al., 2012) and stress (Schwabe, 2017), can reduce prefrontal cortex activity and increase the likelihood of resorting to automated behaviors. It is thus possible that these mental responses have a similar effect on driver behaviors that the cognitive control hypothesis suggests task-directed cognitive control has. In that case, driver support systems should be designed to detect not just high levels of cognitive control, but also high levels of stress and negative emotions, in order to be able to assist drivers in situations where cognitive control is required for safe driving. Other possible mental responses to cognitive tasks, such as emotions (Chan & Singhal, 2015) and fatigue (Ismail & Karwowski, 2020), may have other effects on driver behaviors and thus require other types of driver support.

A model by Dehais et al. (2020) seems useful for understanding which mental responses to variations in task demand might have the greatest impact on traffic safety and thus deserve the most attention. The model suggests that performance in cognitive tasks can be understood in relation to arousal and task engagement (see Figure 3).[4] The authors provide convincing arguments for their model, abandoning the idea of limited resources and instead explain performance variations by referring to observed neurological mechanisms and interactions between brain areas. According to the model, high levels of task engagement and arousal, such as when a driver is highly involved in a challenging and stressful cognitive task, can lead to task perseveration (continuing a behavior even though there is information that strongly speaks against it; Dehais et al., 2010) and inattention blindness/deafness (the inability to perceive clearly visible/audible task-irrelevant stimuli; Rensink, 2009). That is, a

[4] In my interpretation, task engagement is synonymous with task-directed effort and thus typically coincides with cognitive control, while high arousal is synonymous with high levels of emotional stress.

driver might under these conditions "get stuck" performing a cognitive task and fail to notice and/or adapt to system warnings or changes in the traffic environment.

On a neurocognitive level, these phenomena may be (in part) explained by the neurochemicals norepinephrine (associated with arousal; Ross & Van Bockstaele, 2021) and dopamine (associated with, e.g., motivation and attention; Sescousse et al., 2018), which at high concentrations impair executive functioning (Eckstein et al., 2017) and trigger automatic responses (J. R. Wickens et al., 2007). High levels of task engagement also promote the dorsal attention network, which enables focused attention, at the expense of the ventral attention network, which enables the reorienting of attention in response to interruptive stimuli (Vossel et al., 2014). This leads to a deteriorated ability to perceive and adapt to new information (Todd et al., 2005).

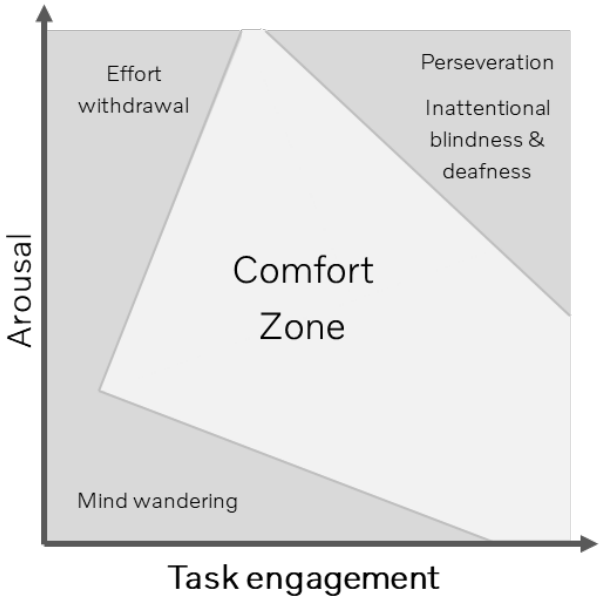

Figure 3. Model of performance degradations depending on arousal and task engagement, adapted from Dehais et al. (2020).

Dehais et al.'s (2020) model further suggests that when task engagement and arousal are low due to non-stimulating and non-rewarding tasks, such as prolonged passive supervisory tasks, performance may suffer due to task disengagement and mind wandering. In this mental state, levels of dopamine and norepinephrine are low, which (just as when they are high) depress executive functioning (Eckstein et al., 2017), thus degrading performance in cognitively controlled activities. Also, during low arousal, the locus coeruleus-norepinephrine system promotes mind-wandering and degrades task performance by fostering activity in task-unrelated brain networks (Chong & Baldwin, 2021).

Moreover, in situations with a high likelihood of task failure, task engagement tends to be low as people give up trying (Fairclough et al., 2019); consequently, task performance degrades due to effort withdrawal (Dehais et al., 2020). Furthermore, in situations where failure can lead to serious consequences, the probability of stress and strong emotional reactions is high at the same time (i.e., arousal is high), which may divert attention away from the task and worsen task performance even more (Dehais et al., 2020).

Worth mentioning, although not included in Dehais et al.'s (2020) model, are fatigue-related performance declines which may occur due to sleepiness or prolonged task engagement. Chong and Baldwin (2021) review neuroergonomic explanations for these deteriorations. They suggest that prolonged execution of cognitively demanding tasks leads to "active fatigue" because the high level of activity in task-activated neurons eventually causes neuronal fatigue. The brain will then engage in sleep-related slow-wave activities to reset the neuronal firing rates in order to achieve cognitive restoration. These sleep-related processes (e.g., local sleep and microsleep, which occur during wakefulness) cause deteriorated task performance and may promote engagement in unrelated tasks, such as mind-wandering (Andrillon et al., 2019). Similarly, these processes also increase with sleepiness caused by long periods spent awake.

Note that unlike the cognitive control hypothesis (Engström et al., 2017), Dehais et al.'s model does not deal with performance in tasks with automated behaviors. Also, their model is primarily intended to explain primary task performance variations in human operators (e.g., pilots), rather than performance degradations during multitasking. Still, it seems appropriate to investigate the effects of different levels of task engagement and arousal, in addition to the level of task-related cognitive control.

In short, there are compelling arguments for questioning the common way of attributing the observed effects of cognitive tasks solely to task-related cognitive control. This simplification is particularly problematic when different studies report different effects, such as the effects of cognitive load on vehicle speed or time headway (see Caird et al., 2018, for a review). Since different mental responses to cognitive demand (e.g., arousal and task engagement) appear to have different effects on driver behaviors, caution needs to be taken when comparing and generalizing study results. Because so many factors (task-related, but also environmental- and human-related) influence which mental responses occur (M. S. Young et al., 2015), different experiments are not necessarily comparable just because they include some sort of cognitive demand. It seems reasonable to assume that participants in controlled experiments who perform artificial tasks under the supervision of test leaders experience a higher level of arousal and task engagement than drivers typically do when talking on a mobile phone during routine driving. Comparing the effects of cell phone use in naturalistic studies with the effects of cognitive tasks in controlled experiments may thus be inappropriate, unless the arousal and engagement are actually comparable. This is not to say that behavioral effects observed in controlled experiments cannot be generalized to actual driving in situations with similar mental responses; however, it is critical to find out which mental responses have a significant influence on driver behaviors and thus need to be kept comparable.

It also follows that drivers might need different types of support during cognitive task engagement, depending on their mental responses (Braun et al., 2021; Dehais et al., 2020). For example, imagine a driver who approaches an intersection while having a casual cell phone conversation with his/her spouse about what to have for dinner. The vehicle's driver monitoring system observes that the driver fails to scan the environment appropriately. A subtle reminder to focus on the traffic environment might be enough to make the driver scan the environment appropriately and notice potential threats. But if the cell phone conversation were instead a heated discussion between spouses in the middle of a divorce, a subtle

reminder would most likely go unnoticed due to the high levels of emotional and cognitive engagement.

6.6 Application areas for physiological measures in traffic safety

The application of physiological measures can facilitate advancement in many different areas within research and product development. The requirements placed on the measures depend, of course, on the application: different measures can be useful for different purposes. To provide concrete examples, a non-exhaustive compilation of application areas related to traffic safety will be described below.

To begin with, physiological measures can be used to validate experimental manipulations by confirming that the intended mental responses indeed occurred. It thus becomes easier to compare the outcomes of different studies applying different experimental manipulations, since there is reason to believe that the mental responses are indeed comparable.

The wind scenario described in Paper IV is an excellent example of the critical importance of being able to confirm the impact of experimental manipulations on mental responses. This scenario was to be used for testing the cognitive control hypothesis (see Nilsson et al., 2017, Section 3.3, for a brief description of the experiment). Lane keeping is a consistent and extensively practiced task for experienced drivers, so it can be assumed to be automatized and should thus not suffer from performance degradations during cognitive load, according to the cognitive control hypothesis (Engström et al., 2017). But, if lane keeping is made sufficiently difficult, it should require cognitive control and consequently be negatively affected by cognitive load (Engström et al., 2017). This effect was observed in a previous study by Medeiros-Ward et al. (2014), and the wind scenario in Paper IV was set up to replicate that study. Contrary to what was expected, lane keeping was found to improve when the drivers were engaged in a cognitive task, with and without the crosswind (Nilsson et al., 2017). The results thus seem to contradict the cognitive control hypothesis. However, an analysis of the physiological data showed that the crosswind had no statistically significant effect on any of the physiological measures that were derived (including six measures that were sensitive to other cognitive load manipulations in the same experiment). The unchanged physiological measures strongly indicate that the crosswind did not actually have any substantial effect on the drivers' mental state. Thus the goal of the experimental crosswind manipulations—to make lane keeping cognitively demanding—was clearly not reached. The results could therefore not be used to confirm or reject the cognitive control hypothesis.

Including multiple measures from different measurement categories (i.e., physiological, self-report, and performance) improves the chances of being able to differentiate between mental responses further (Hancock & Matthews, 2019), providing a more detailed understanding of a situation (as pointed out in Section 6.4). As an example, such improved understanding comes from interpreting Paper IV's results together with those of Paper II, which revealed that the drivers had fewer glances towards potential threats in off-path locations when they were engaged in a cognitive task. Clearly, directing the gaze to relevant areas in the traffic environment is a prerequisite for safe driving. However, gaze direction alone does not guarantee that drivers perceive and understand the fixated information, nor that they will act appropriately in response (Brown, 2005; Strayer et al., 2004). At this point, the physiological

data from Paper IV can add to our understanding. An increase in pupil diameter, together with a decrease in blink rate and blink duration, indicates that the drivers increased their visual attention and effort when approaching and passing the potential threats in all load conditions. That is, the physiological data suggest the driver did indeed perform attentive glances. It does, however, appear that the amount of extra effort that the drivers invested when the driving demand increased, decreased with higher levels of cognitive task demand, as the pupil diameter responses decreased in magnitude (see Figure 2, Paper IV). To avoid excessive testing—and increased risk of Type I errors—this decrease in pupil diameter increase was not statistically tested and thus needs to be validated in future experiments. Also observable in Figure 2 (Paper IV) but not statistically tested is the temporary drop in heart rate as the drivers approached the critical scenarios. One might speculate that this response reflects a preparation for action (Cooke, 2013). As with the pupil diameter and eye blink responses, this drop in heart rate decreased when the cognitive task demand increased. Again, further research is needed to replicate and statistically test this finding, as well as to confirm this interpretation. In short, the drivers directed their gaze to relevant off-path locations, increased their effort and visual attention (meaning that the glances were attentive), and may even have prepared to act (indicating that they understood the situation). When cognitive task demand increased, they still adapted to the scenario, but seem to have done so to a lesser extent.

When more detailed analyses and mental state assessments can be performed, it follows that our understanding of the mechanisms underlying performance degradations will increase. For example, measurements of ERPs (recall that ERPs are brain responses to stimuli) can show the effects of experimental manipulations at specific stimulus-processing stages (Joos et al., 2014). Van der Heiden et al. (2021), for example, showed that during AD, drivers had smaller responses in the frontal P3 component (reflecting a subject's ability to notice stimuli) to task-irrelevant auditory stimuli when performing cognitive tasks. Based on this finding, they inferred that engagement in cognitive tasks during AD could cause delayed or absent responses to auditory alerts.

In addition to providing a more detailed understanding of mental responses to task demands or situations, physiological measures are also useful in mental state assessments in studies requiring high levels of ecological validity. The reason for this is that physiological measures—provided they are unintrusively recorded—can offer mental state assessments without altering the driving task (unlike, e.g., the DRT and LCT) or interrupting the driver (unlike, e.g., self-reports). The physiological measures therefore allow more life-like study designs than most other mental state assessment techniques.

Finally, physiological measures could also be used in advanced driver assistance systems to monitor drivers' mental or physical state so that the right support can be provided at the right time to prevent safety-critical performance degradations without nuisance interventions. The measures can, for example, be used in the assessment of a driver's ability to take over control after AD, so that the system can decide if the car should make a safe stop or hand over control (Weaver & DeLucia, 2020), and what such a hand-over should look like (see e.g. Niu & Ma, 2022, on how a fatigue warning before a take-over request may improve driver performance). They may also be used to detect a sudden illness that requires the car to take over control from the driver and perform a safe stop (Langer et al., 2016). In-vehicle driver state monitoring systems have to perform highly accurate, real-time mental and/or physical

state assessments using only fully non-intrusive recording techniques. This is certainly not easy, but plenty of effort in this area is already being made (for reviews, see Arakawa, 2021; Leonhardt et al., 2018; Sidikova et al., 2020).

6.7 Future work

The findings reported in this thesis need to be replicated and validated in future studies using, for example, other test participant criteria and test environments. Also, individual differences in how cognitive task engagement affects behaviors and abilities deserve greater investigation. That is, are there certain individual characteristics, such as older age or a risk-taking personality, which entail an increased risk of driving performance degradations during cognitive task engagement? If so, higher-risk groups should be identified, since they deserve more attention and representation in product development and evaluation.

As argued repeatedly in this thesis, future studies of cognitive load are likely to provide the most useful insights if researchers acknowledge the existence of multiple mental responses to cognitive task demands. Study results can then help them determine which mental responses have safety-relevant effects on driver behaviors and abilities and explore the neurocognitive mechanisms that underly changes in behavior. At the same time, the level of detail and precision in driver state assessments needs to be balanced against measurability and usability in the context where the research is taking place. Some suggestions about what mental states should be measured to predict and counteract performance degradations in car drivers have already been put forward by others. For example, Lohani et al. (2019) argue that in a real-world setting, it should be sufficient to classify the drivers' mental state along an arousal spectrum that ranges from passive states with low levels of arousal (e.g., fatigue), to highly aroused states, including high levels of cognitive load. Dehais et al. (2020), on the other hand, argue that this spectrum is not enough to understand performance degradations; one also needs a dimension of task engagement (as discussed in Section 6.5). Others argue that one cannot ignore the drivers' emotional responses to task engagement (and other influencing factors), since they also affect driver behaviors and traffic safety (Braun et al., 2021). It can also be argued that the use of cognitive control deserves monitoring, since performance degradations can then be predicted based on the cognitive control hypothesis (Engström et al., 2017). Future work should thus continue to disentangle the various mental responses to cognitive task demand so that their individual and combined effects on driver behaviors can be understood. Only then can we know which mental states that should be measured in order to detect and counteract performance degradations.

Another area of research which deserves investigation is drivers' ability to shift attention (Lee, 2014). That is, how easy it is for the driver to disengage from some non-driving-related cognitive activity and re-engage in the driving task when needed? The activity could be an internally driven, non-driving-related cognitive activity such as mind wandering or thinking—or an externally driven NDRT such as conversing or interacting with some device. The ability to shift attention and re-engage in driving is critical after AD when the driver may have been disengaged from the driving task for quite some time. (Research efforts to understand this ability are plenty; see Weaver & DeLucia, 2020, for a review.) This ability is, however, also critical, but rarely investigated, in manual and assisted driving when the driver

multitasks by sharing and shifting attention between the driving task and some non-driving-related activity (Lee, 2014). If drivers are able to let go of the non-driving-related activity and shift attention to the driving task whenever needed, the diversion of attention towards the non-driving-related activity does not pose a problem. Future studies should therefore investigate what factors influence drivers' task-switching abilities in manual driving, assisted driving, and AD. In addition to the factors that affect drivers' ability to take over control after AD, factors that determine how drivers time-share and prioritize between tasks (depending on the traffic situation, type of task, and mental state) should also be investigated. As pointed out by Dehais et al. (2020) and discussed in Section 6.5, high levels of arousal and task engagement may lead to task perseveration—but other factors may as well (Fox & Huffman, 2002). For example, the chance that drivers exhibit long off-path glances when engaging in a visually demanding NDRT has been found to increase as a function of task duration, as the drivers' desire to complete the task grows with time and task perseveration thus increases (Lee et al., 2012).

Finally, because of the many advantages of using physiological measures for assessing driver mental states, the measures should be used to a greater extent in future studies whenever drivers' mental state plays a role. Further research is still needed to establish the best measures and methods for reliably assessing drivers' mental states—ideally, in real time. Additionally, efforts should be made to understand and deal with individual differences since the variability between individuals is large (Stemmler & Wacker, 2010). Research should determine which individual differences in physiological responses are due to differences between individuals' mental responses to the situation at hand (e.g., some participants get more stressed by being asked to perform a cognitive task than others; Matthews & Campbell, 2009; Szalma, 2009) and which are due to differences between individuals' physiological responses to the same mental response (e.g., frontal theta power increases during cognitive engagement in most people, but not all; Mitchell et al., 2008). Due to individual differences in physiological responses, individually adapted algorithms for reliable driver state assessments should be developed (Noh et al., 2021). Finally, future work should also explore the feasibility of continual learning— that is, systems which continuously learn the driver's physiological response patterns over time, in order to recognize typical and atypical driver states for that individual.

7 Conclusions

In this thesis, cognitive load in car drivers was studied from four different safety-related perspectives: investigating the effects of cognitive tasks on behaviors and decisions in three different stages of traveling (in a critical event [Paper I], in routine driving [Paper II], and in the pre-trip stage where decisions about how to travel are being made [Papers III]) and exploring how cognitive load in car drivers can be assessed (Paper IV). The main findings related to the four research questions and their implications are summarized below.

1. How do cognitive tasks affect drivers' response times in an unexpected safety-critical event?

In a safety-critical event with a lead vehicle braking unexpectedly, drivers' engagement in a cognitive task did not affect their response times (Paper I). This finding means that it is inappropriate to generalize from the (albeit well-established and highly consistent) effect of delayed response times in artificial response tasks, such as the DRT, to more realistic safety-critical events in which responses are triggered automatically by visual looming.

2. How do cognitive tasks affect drivers' glance behavior in routine driving?

During routine driving in traffic environments with potential threats in off-path locations, cognitive tasks affected drivers' glance behaviors by inducing gaze concentration (Paper II), in line with previous research. Novel with this study was the high temporal resolution used to explore how the gaze concentration phenomenon related to safety-relevant elements in the traffic environment and to variations in cognitive load. While gaze shifts from an on-path to an off-path direction were inhibited by cognitive load, gaze shifts in the opposite direction (off-path to on-path) were unaffected. Further, the drivers' visual adaptation to the traffic scenarios in terms of the timing of glances towards situation-relevant off-path locations was unaffected by cognitive load; only the total number of off-path glances was reduced. That is, gaze shifts were cancelled rather than delayed. This finding suggests that gaze direction control is not affected by cognitive load; however, gaze shift initiations are instantaneous and transient, and gaze shifts are permanently omitted if not initiated in time.

The cognitive control hypothesis (Engström et al., 2017) offers a plausible explanation for the observed effects of cognitive load on both response times and gaze shifts. The hypothesis stipulates that cognitive load only affects behaviors which require cognitive control (i.e., less practiced or inconsistent behaviors, such as looking away from the road ahead or pressing a button in response to an artificial stimulus) but leaves automatized behaviors (i.e., consistent and extensively practiced behaviors, such as braking in response to visual looming or directing the gaze towards the road ahead) unaffected. Consequently, the hypothesis suggests that cognitive tasks are detrimental to traffic safety only when there is no suitable automatized behavior for the driver to fall back on (Engström et al., 2017). The cognitive control hypothesis thus seems useful for predicting when drivers' engagement in cognitive tasks may compromise safety—and where driver support is thus needed.

3. To what extent are drivers able to engage in non-driving-related tasks while traveling in AD mode in real traffic?

Drivers were able to engage in a non-driving-related task which was cognitively, visually, and motorically demanding, and perform it well, while traveling in AD mode in real traffic (Paper III). This finding suggests that drivers' desire to engage in non-driving-related activities while traveling using AD can be fulfilled. Being able to work or engage in other non-driving-related activities may influence drivers' decisions about how to travel by motivating them to invest in and use AD systems—systems which are expected to increase traffic safety. In this way, cognitive tasks could change from being a safety problem to a safety catalyst by increasing the use of AD systems—given that AD is indeed safer than manual driving.

4. How can cognitive load in car drivers be assessed using physiological measures?

The results in Paper IV demonstrate that for accurate assessments of cognitive load, it has to be acknowledged that cognitive load consists of multiple mental responses which can give rise to different physiological responses. To date, attempts to assess cognitive load normally employ a unidimensional approach; that is, they only consider the level, not the composition (i.e., the combination of mental responses), of cognitive load. This approach entails a risk of deriving measures which are only sensitive to certain mental responses—responses which may or may not be part of the cognitive load in a different context. Such measures might thus not be useful for detecting cognitive load in contextually different settings. Concurrent analyses of multiple mental responses using multiple physiological measures can, however, greatly improve the assessment of drivers' mental states and increase the usability of physiological measures in research and product development.

Acknowledging multiple mental responses to cognitive task demands might also help resolve some inconsistencies between studies, in which cognitive tasks were found to have different effects on both physiological responses and driver behaviors in different studies. While some effects of cognitive tasks appear to be universal, in the sense that they are common for most cognitive tasks, many others differ depending on the task and the situation as a whole. It follows that efforts should be made to understand which mental responses to cognitive task engagement have the greatest impact on driver behaviors related to traffic safety, preferably by considering the neurocognitive mechanisms that underlie the effects. Such knowledge, together with reliable measures of the relevant mental responses, enables the development of systems that provide drivers with the right support at the right time and which can thus improve traffic safety.

References

Abd Rahman, N. I., Md Dawal, S. Z., & Yusoff, N. (2020). Driving mental workload and performance of ageing drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *69*, 265–285. https://doi.org/10.1016/j.trf.2020.01.019

Ahlström, C., Solis-marcos, I., Nilsson, E., & Åkerstedt, T. (2020). The impact of driver sleepiness on fixation-related brain potentials. *Journal of Sleep Research, 29*, Article e12962. https://doi.org/10.1111/jsr.12962

Ahlström, C., Victor, T., Wege, C., & Steinmetz, E. (2012). Processing of eye/head-tracking data in large-scale naturalistic driving data sets. *IEEE Transactions on Intelligent Transportation Systems*, *13*(2), 553–564. https://doi.org/10.1109/TITS.2011.2174786

Alm, H., & Nilsson, L. (1995). The effects of a mobile telephone task on driver behaviour in a car following situation. *Accident Analysis and Prevention*, *27*(5), 707–715. https://doi.org/10.1016/0001-4575(95)00026-V

Americal Automobile Association. (2019). *Automated Vehicle Survey — Phase IV*. https://www.infrastructureusa.org/wp-content/uploads/2019/03/AV-Consumer-Survey-Fact-Sheet-FINAL-2-28-19.pdf

Andersson, J., Asker, E., Habibovic, A., Klingegård, M., Lindström, D., & Malmsten Lundgren, V. (2018). *Method development for studies at test tracks of long-term behavioral adaptation in automated vehicles* (Project report A-0018). https://www.saferresearch.com/library/final-report-0018

Andrillon, T., Windt, J., Silk, T., Drummond, S. P. A., Bellgrove, M. A., & Tsuchiya, N. (2019). Does the Mind Wander When the Brain Takes a Break? Local Sleep in Wakefulness, Attentional Lapses and Mind-Wandering. *Frontiers in Neuroscience*, *13,* Article 949. https://doi.org/10.3389/fnins.2019.00949

Angell, L., Auflick, J., Austria, P. A., Kochhar, D., Tijerina, L., Biever, W., Diptman, T., Hogsett, J., & Kiger, S. (2006). *Driver Workload Metrics Project Task 2 Final Report* (Report No. DOT HS 810 635), National Highway Traffic Safety Administration. https://trid.trb.org/view/855927

Annett, J. (2002). Subjective rating scales: Science or art? *Ergonomics*, *45*(14), 966–987. https://doi.org/10.1080/00140130210166951

Arakawa, T. (2021). A review of heartbeat detection systems for automotive applications. *Sensors*, *21*(18), Article 6112. https://doi.org/10.3390/s21186112

Baker, J. M., Rojas-Valverde, D., Gutiérrez, R., Winkler, M., Fuhrimann, S., Eskenazi, B., Reiss, A. L., & Mora, A. M. (2017). Portable functional neuroimaging as an environmental epidemiology tool: A how-to guide for the use of fNIRS in field studies. *Environmental Health Perspectives*, *125*(9), Article 094502. https://doi.org/10.1289/EHP2049

Balters, S., & Steinert, M. (2015). Capturing emotion reactivity through physiology measurement as a foundation for affective engineering in engineering design science and engineering practices. *Journal of Intelligent Manufacturing, 28, 1585-1607*. https://doi.org/10.1007/s10845-015-1145-2

Bari, D. S., Aldosky, H. Y. Y., Tronstad, C., Kalvøy, H., & Martinsen, G. (2018). Electrodermal responses to discrete stimuli measured by skin conductance, skin potential, and skin susceptance. *Skin Research and Technology*, *24*(1), 108–116. https://doi.org/10.1111/srt.12397

Barkley, R. A. (2012). *Executive functions: what they are, how they work, and why they evolved*. Guilford Publications. https://doi.org/10.5860/choice.50-2366

Barzegaran, E., Vildavski, V. Y., & Knyazeva, M. G. (2017). Fine Structure of Posterior Alpha Rhythm in Human EEG: Frequency Components, Their Cortical Sources, and Temporal Behavior. *Scientific Reports*, *7,* Article 8249. https://doi.org/10.1038/s41598-017-08421-z

Baumann, M., Petzoldt, T., Groenewoud, C., Hogema, J., & Krems, J. F. (2008). The Effect of Cognitive Tasks on Predicting Events in Traffic. *Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems*, 3–12.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265. https://doi.org/10.1037/0022-3514.74.5.1252

Belyusar, D., Mehler, B., Solovey, E., & Reimer, B. (2015). Impact of repeated exposure to a multilevel working memory task on physiological arousal and driving performance. *Transportation Research Record: Journal of the Transportation Research Board*, *2518*(1), 46–53. https://doi.org/10.3141/2518-06

Benarroch, E. E. (2011). Central Autonomic Control. In P. A. Low, R. J. Polinsky, D. Robertson, & J. F. R. Paton (Eds.), *Primer on the Autonomic Nervous System* (3rd ed.). Elsevier Science & Technology.

Billman, G. E. (2011). Heart rate variability - A historical perspective. *Frontiers in Physiology*, *2,* Article 86. https://doi.org/10.3389/fphys.2011.00086

Billman, G. E. (2013). The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. *Frontiers in Physiology*, *4*, Article 26. https://doi.org/10.3389/fphys.2013.00026

Biondi, F., Balasingam, B., & Ayare, P. (2021). On the Cost of Detection Response Task Performance on Cognitive Load. *Human Factors*, *63*(5), 804–812. https://doi.org/10.1177/0018720820931628

Biondi, F., Turrill, J. M., Coleman, J. R., Cooper, J. M., & Strayer, D. L. (2015). Cognitive Distraction Impairs Drivers' Anticipatory Glances: An On-Road Study. *Proceedings of the Eighth International Driving Sypmosium on Human Factors in Driver Assessment, Training and Vehicle Design*, 23–29.

Bjorvatn, A., Page, Y., Fahrenkrog, F., Weber, H., Aittoniemi, E., Heum, P., Lehtonen, E., Silla, A., Bärgman, J., & Al., E. (2021). *Impact Evaluation Results* (L3Pilot Deliverable D7.4). https://l3pilot.eu/fileadmin/user_upload/Downloads/Deliverables/Update_14102021/L3Pilot-SP7-D7.4-Impact_Evaluation_Results-v1.0-for_website.pdf

Boda, C. N., Dozza, M., Bohman, K., Thalya, P., Larsson, A., & Lubbe, N. (2018). Modelling how drivers respond to a bicyclist crossing their path at an intersection: How do test track and driving simulator compare? *Accident Analysis and Prevention*, *111*, 238–250. https://doi.org/10.1016/j.aap.2017.11.032

Botvinick, M. M., & Cohen, J. D. (2014). The computational and neural basis of cognitive control: Charted territory and new frontiers. *Cognitive Science*, *38*(6), 1249–1285. https://doi.org/10.1111/cogs.12126

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*(8), 1017–1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x

Braun, M., Weber, F., & Alt, F. (2021). Affective Automotive User Interfaces -- Reviewing the State of Emotion Regulation in the Car. *ACM Computing Surveys*, *54*(7), Article 137. http://arxiv.org/abs/2003.13731

Broström, R., Rydström, A., & Kopp, C. (2019). Drivers Quickly Trust Autonomous Cars. In T. Z. Ahram & C. Falcão (Eds.), *Advances in Usability, User Experience and Assistive Technology. AHFE 2018. Advances in Intelligent Systems and Computing* (Vol. 794). Springer, Cham. https://doi.org/https://doi.org/10.1007/978-3-319-94947-5_69

Brouwer, A. M., Hogervorst, M. A., Holewijn, M., & van Erp, J. B. F. (2014). Evidence for effects of task difficulty but not learning on neurophysiological variables associated with effort. *International Journal of Psychophysiology*, *93*(2), 242–252. https://doi.org/10.1016/j.ijpsycho.2014.05.004

Brown, I. (2005). *Review of the ' looked but failed to see ' accident causation factor* (Road Safety Research Report No . 60). Department for Transport, England.

Bruyas, M.-P., & Dumont, L. (2013). Sensitivity of Detection Response Task (DRT) to the Driving Demand and Task Difficulty. *PROCEEDINGS of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, 64–70.

Bruyas, M.-P., Brusque, C., Tattegrain, H., Auriault, A., Aillerie, I., & Duraz, M. (2008). Consistency and sensitivity of lane change test according to driving simulator characteristics. *IET Intelligent Transport Systems*, *2*(4), 306–314. https://doi.org/10.1049/iet-its:20080041

Burke, D. A., Rotstein, H. G., & Alvarez, V. A. (2017). Striatal Local Circuitry: A New Framework for Lateral Inhibition. *Neuron*, *96*(2), 267–284. https://doi.org/10.1016/j.neuron.2017.09.019

Burle, B., Spieser, L., Roger, C., Casini, L., Hasbroucq, T., & Vidal, F. (2015). Spatial and temporal resolutions of EEG: Is it really black and white? A scalp current density view. *International Journal of Psychophysiology*, *97*(3), 210–220. https://doi.org/10.1016/j.ijpsycho.2015.05.004

Cabrall, C. D. D., Eriksson, A., Dreger, F., Happee, R., & de Winter, J. (2019). How to keep drivers engaged while supervising driving automation? A literature survey and categorisation of six solution areas. *Theoretical Issues in Ergonomics Science*, *20*(3), 332–365. https://doi.org/10.1080/1463922X.2018.1528484

Caird, J. K., Johnston, K. A., Willness, C. R., Asbridge, M., & Steel, P. (2014). A meta-analysis of the effects of texting on driving. *Accident Analysis and Prevention*, *71*, 311–318. https://doi.org/10.1016/j.aap.2014.06.005

Caird, J. K., Simmons, S. M., Wiley, K., Johnston, K. A., & Horrey, W. J. (2018). Does Talking on a Cell Phone, With a Passenger, or Dialing Affect Driving Performance? An Updated Systematic Review and Meta-Analysis of Experimental Studies. *Human Factors*, *60*(1), 101–133. https://doi.org/10.1177/0018720817748145

Cardona, G., García, C., Serés, C., Vilaseca, M., & Gispets, J. (2011). Blink rate, blink amplitude, and tear film integrity during dynamic visual display terminal tasks. *Current Eye Research*, *36*(3), 190–197. https://doi.org/10.3109/02713683.2010.544442

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: Meta-analytic evidence. *Journal of Physiology - Paris*, *109*(1-3), 3–15. https://doi.org/10.1016/j.jphysparis.2014.04.003

Chan, M., & Singhal, A. (2015). Emotion matters: Implications for distracted driving. *Safety Science*, *72*, 302–309. https://doi.org/10.1016/j.ssci.2014.10.002

Chen, F., Zhou, J., Wang, Y., Yu, K., Arshad, S. Z., Khawaji, A., & Conway, D. (2016). *Robust multimodal cognitive load measurement* (D. Tan & J. Vanderdonckt (Eds.)). Springer.

Chihara, T., Kobayashi, F., & Sakamoto, J. (2020). Evaluation of mental workload during automobile driving using one-class support vector machine with eye movement data. *Applied Ergonomics*, *89*, Article 103201. https://doi.org/10.1016/j.apergo.2020.103201

Choi, J. W., & Kim, K. H. (2018). *Computational EEG Analysis* (C.-H. Im (Ed.)). Springer. https://doi.org/10.1007/978-981-13-0908-3

Chong, S. D., & Baldwin, C. L. (2021). The Origins of Passive, Active, and Sleep-Related Fatigue. *Frontiers in Neuroergonomics*, *2*, Article 765322. https://doi.org/10.3389/fnrgo.2021.765322

Chuang, L. L., Donker, S. F., Kun, A. L., & Janssen, C. P. (2018). Workshop on the mobile office. Adjunct Proceedings - 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2018, 10–16. https://doi.org/10.1145/3239092.3239094

Clayton, M. S., Yeung, N., & Cohen Kadosh, R. (2015). The roles of cortical oscillations in sustained attention. *Trends in Cognitive Sciences*, *19*(4), 188–195. https://doi.org/10.1016/j.tics.2015.02.004

Cohen, J. D. (2017). Cognitive Control - Core Constructs and Current Considerations. In T. Egner (Ed.), *The Wiley Handbook of Cognitive Control* (1st ed., pp. 3–28). John Wiley & Sons, Ltd.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the stroop effect. *Psychological Review*, *97*(3), 332–361. https://doi.org/10.1037/0033-295X.97.3.332

Collet, C., Guillot, A., & Petit, C. (2010). Phoning while driving I: A review of epidemiological, psychological, behavioural and physiological studies. *Ergonomics*, *53*(5), 589–601. https://doi.org/10.1080/00140131003672023

Collet, C., & Musicant, O. (2019). Associating vehicles automation with drivers functional state assessment systems: A challenge for road safety in the future. *Frontiers in Human Neuroscience*, *13*, Article 131. https://doi.org/10.3389/fnhum.2019.00131

Conti, A. S., Dlugosch, C., Vilimek, R., Keinath, A., & Bengler, K. (2012). An assessment of cognitive workload using detection response tasks. In N. A. Stanton (Ed.), *Advances in Human Aspects of Road and Rail Transportation* (pp. 735–743). CRC Press, Taylor & Francis Group. https://doi.org/10.1201/b12320

Conway, D., Dick, I., Li, Z., Wang, Y., & Chen, F. (2013). The effect of stress on cognitive load measurement. In P. Kotzé, G. Marsden, G. Lindgaard, J. Wesson, & M. Winckler (Eds.), *Human-Computer Interaction - INTERACT 2013, Lecture notes in computer science* (Vol. 8120, pp. 659–666). Springer, Berlin, Heidelberg. https://doi.org/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. *International Review of Sport and Exercise Psychology*, *6*(1), 122–138. https://doi.org/10.1080/1750984X.2012.724438

Cori, J. M., Anderson, C., Shekari Soleimanloo, S., Jackson, M. L., & Howard, M. E. (2019). Narrative review: Do spontaneous eye blink parameters provide a useful assessment of state

drowsiness? *Sleep Medicine Reviews*, *45*, 95–104. https://doi.org/https://doi.org/10.1016/j.smrv.2019.03.004

Daly, I., Pichiorri, F., Faller, J., Kaiser, V., Kreilinger, A., Scherer, R., & Muller-Putz, G. (2012). What does clean EEG look like? *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 3963–3966. https://doi.org/10.1109/EMBC.2012.6346834

Dampney, R. A. L. (2019). Central mechanisms generating cardiovascular and respiratory responses to emotional stress. In G. Fink (Ed.), *Stress: Physiology, Biochemistry, and Pathology. Handbook of Stress Series, Volume 3* (pp. 391–402). Academic Press. https://doi.org/10.1016/B978-0-12-813146-6.00028-X

Dawson, M. E., Schell, A. M., & Filion, D. L. (2016). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology, Fourth Edition* (pp. 217–243). Cambridge University Press. https://doi.org/10.1017/9781107415782.010

de Waard, D. (1996). *The Measurement of Drivers' Mental Workload* [Doctoral dissertation, University of Groningen]. http://apps.usd.edu/coglab/schieber/pdf/deWaard-Thesis.pdf

de Winter, J. C. F., van Leeuwen, P. M., & Happee, R. (2012). Advantages and Disadvantages of Driving Simulators: A Discussion. In A. J. Spink, O. E. Grieco, L. W. S. Krips, L. W. S. Loijens, L. P. J. J. Noldus, & P. H. Zimmerman (Eds.), *Proceedings of the Measuring Behavior Conference* (pp. 47–50).

de Winter, J. C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 196–217. https://doi.org/10.1016/j.trf.2014.06.016

Dehais, F., Hodgetts, H. M., Causse, M., Behrend, J., Durantin, G., & Tremblay, S. (2019). Momentary lapse of control: A cognitive continuum approach to understanding and mitigating perseveration in human error. *Neuroscience and Biobehavioral Reviews*, *100,* 252–262. https://doi.org/10.1016/j.neubiorev.2019.03.006

Dehais, F., Lafont, A., Roy, R., & Fairclough, S. (2020). A Neuroergonomics Approach to Mental Workload, Engagement and Human Performance. *Frontiers in Neuroscience*, *14*, Article 268. https://doi.org/10.3389/fnins.2020.00268

Dehais, F., Tessier, C., Christophe, L., & Reuzeau, F. (2010). The Perseveration Syndrome in the Pilot's Activity: Guidelines and Cognitive Countermeasures. In P. Palanquey, J. Vanderdonckt, & M. Winckler (Eds.), *Human Error, Safety and Systems Development. Lecture Notes in Computer Science* (Vol. 5962, pp. 68–80). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-11750-3_6

Deme, D. (2019). Review on Factors Causes Road Traffic Accident In Africa. *Journal of Civil Engineering Research & Technology*, *1*(1), 1-8. https://doi.org/10.47363/jcert/2019(1)101

Desmet, C., & Diependaele, K. (2019). An eye-tracking study on the road examining the effects of handsfree phoning on visual attention. *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*, 549–559. https://doi.org/10.1016/j.trf.2018.11.013

Di Flumeri, G., Borghini, G., Aricò, P., Sciaraffa, N., Lanzi, P., Pozzi, S., Vignali, V., Lantieri, C., Bichicchi, A., Simone, A., & Babiloni, F. (2018). EEG-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings. *Frontiers in Human Neuroscience*, *12*, Article 509. https://doi.org/10.3389/fnhum.2018.00509

Diels, C., Robbins, R., & Reed, N. (2011). Behavioural validation of the TRL driving simulator digicar: Phase 1 - Speed choice. *Driver Behaviour and Training*, *5*, 429–445.

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, *113*(10), 2636–2641. https://doi.org/10.1073/pnas.1513271113

Dingus, T. A., Owens, J. M., Guo, F., Fang, Y., Perez, M., McClafferty, J., Buchanan-King, M., & Fitch, G. M. (2019). The prevalence of and crash risk associated with primarily cognitive secondary tasks. *Safety Science*, *119*, 98–105. https://doi.org/10.1016/j.ssci.2019.01.005

Dolan, R. J. (2002). Emotion, Cognition, and Behavior. *Science*, *298*(5596), 1191–1194. https://doi.org/10.1126/science.1076358

Drews, F. A., Yazdani, H., Godfrey, C. N., Cooper, J. M., & Strayer, D. L. (2009). Text messaging during simulated driving. *Human Factors*, *51*(5), 762–770. https://doi.org/10.1177/0018720809353319

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*(10), 475–483. https://doi.org/10.1016/s0166-2236(00)01633-7

Dunn, N., Dingus, T., & Soccolich, S. (2019). Understanding the Impact of Technology: Do Advanced Driver Assistance and Semi-Automated Vehicle Systems Lead to Improper Driving Behavior? (Technical Report). AAA Foundation for Traffic Safety, Washington, D.C. https://aaafoundation.org/wp-content/uploads/2019/12/19-0460_AAAFTS_VTTI-ADAS-Driver-Behavior-Report_Final-Report.pdf

Dussault, C., Jouanin, J. C., Philippe, M., & Guezennec, C. Y. (2005). EEG and ECG changes during simulator operation reflect mental workload and vigilance. *Aviation Space and Environmental Medicine*, *76*(4), 344–351.

Eckstein, M. K., Guerra-Carrillo, B., Miller Singley, A. T., & Bunge, S. A. (2017). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive Neuroscience*, *25*, 69–91. https://doi.org/10.1016/j.dcn.2016.11.001

Eggemeier, F. T., Wilson, G. F., Kramer, A. F., & Damos, D. L. (1991). Workload assessment in multi-task environments. In D. Damos (Ed.), *Multiple-task performance* (pp. 207–216). CRC Press. https://doi.org/https://doi.org/10.1201/9781003069447

Elvik, R., Meyer, S. F., Hu, B., Ralbovsky, M., Vorwagner, A., & Boghani, H. (2020). *Methods for forecasting the impacts of connected and automated vehicles* (Deliverable D3.2 of the H2020 project LEVITATE). https://levitate-project.eu/wp-content/uploads/2020/08/D3.2-Methods-for-forecasting-the-impacts-of-connected-and-automated-vehicles-Final.pdf

Engström, J. (2010). Scenario criticality determines the effects of working memory load on brake response time. In J. Krems, T. Petzoldt, & M. Henning (Eds.), *Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems* (pp. 25–36). Lyon, France.

Engström, J., Aust, M. L., & Viström, M. (2010). Effects of working memory load and repeated scenario exposure on emergency braking performance. *Human Factors*, *52*(5), 551–559. https://doi.org/10.1177/0018720810381072

Engström, J., Larsson, P., & Larsson, C. (2013). Comparison of Static and Driving Simulator Venues for the Tactile Detection Response Task. *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* (pp. 369–375).

Engström, J., Markkula, G., Victor, T., & Merat, N. (2017). Effects of Cognitive Load on Driving Performance: The Cognitive Control Hypothesis. *Human Factors*, *59*(5), 734–764. https://doi.org/10.1177/0018720817690639

Engström, J., Monk, C., Hanowski, R., Horrey, W., Lee, J., McGehee, D., Regan, M., Stevens, A., Traube, E., Tuukkanen, M., Victor, T., & Yang, D. (2013). *A Conceptual Framework and Taxonomy for Understanding and Categorizing Driver Inattention*. US-EU ITS Cooperation, Driver Distraction and HMI Working Group.

Engström, J., Victor, T. & Markkula, G. (2013). Attention Selection and Multitasking in Everyday Driving: A Conceptual Model. In M. Reagan, J. D. Lee, & T. W. Victor (Ed.), *Driver Distraction and Inattention: Advances in Research and Countermeasures* (pp. 27–54). ASHGATE.

Eslinger, P. J. (1996). Conceptualizing, describing, and measuring components of executive function. In G. R. Lyon & N. A. Kransnegor (Eds.), *Attention, memory, and executive function* (pp. 367–395).

Eugensson, A., Ivarsson, J., Lie, A., & Tingvall, C. (2011). Cars are driven on roads, joint visions and modern technologies stress the need for co-operation. *Paper Presented at the 22nd International Technical Conference on the Enhanced Safety of Vehicles,* Washington, D.C.

Fairclough, S., Ewing, K., Burns, C., & Kreplin, U. (2019). Neural efficiency and mental workload: Locating the red line. In A. Johnson & R. W. Proctor (Eds.), *Neuroergonomics: The Brain at Work and in Everyday Life* (pp. 73–77). Academic Press. https://doi.org/10.1016/B978-0-12-811926-6.00012-9

Fairclough, S., & Mulder, L. J. M. (2012). Psychophysiological processes of mental effort investment. In R. A. Wright & H. A. Gendolla (Eds.), *How motivation affects cardiovascular response: Mechanisms and applications* (pp. 61–76). American Psychological Association. https://doi.org/https://doi.org/10.1037/13090-003

Faure, V., Lobjois, R., & Benguigui, N. (2016). The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, *40*, 78–90. https://doi.org/10.1016/j.trf.2016.04.007

Fisher, D. L. (2015). Cognitive Distraction in the Wild: Next Steps-Addressing a Not-So-Humdrum Conundrum. *Human Factors*, *57*(8), 1331–1333. https://doi.org/10.1177/0018720815576435

Fitch, G. M., Soccolich, S. a, Guo, F., McClafferty, J., Fang, Y., Olson, R. L., Perez, M. A., Hanowski, R. J., Hankey, J. M., & Dingus, T. A. (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). U.S. Department of Transportation, National Highway Traffic Safety Administration.

Fors, C., Ahlstrom, C., & Anund, A. (2013). Simulator validation with respect to driver sleepiness and subjective experiences: Final report of the project SleepEYE II, part 1 (ViP Publication 2013-1).

Fox, S., & Huffman, M. (2002). Escalation behavior as a specific case of goal-directed activity: A persistence paradigm. *Basic and Applied Social Psychology*, *24*(4), 273–285. https://doi.org/10.1207/S15324834BASP2404_3

Foy, H. J., & Chapman, P. (2018). Mental workload is reflected in driver behaviour, physiology, eye movements and prefrontal cortex activation. *Applied Ergonomics*, *73*, 90–99. https://doi.org/10.1016/j.apergo.2018.06.006

Gärtner, M., Grimm, S., & Bajbouj, M. (2015). Frontal midline theta oscillations during mental arithmetic: Effects of stress. *Frontiers in Behavioral Neuroscience*, *9*, Article 96. https://doi.org/10.3389/fnbeh.2015.00096

Goschke, T., & Bolte, A. (2014). Emotional modulation of control dilemmas: The role of positive affect, reward, and dopamine in cognitive stability and flexibility. *Neuropsychologia*, *62*, 403–423. https://doi.org/10.1016/j.neuropsychologia.2014.07.015

Grassmann, M., Vlemincx, E., Von Leupoldt, A., Mittelstädt, J. M., & Van Den Bergh, O. (2016). Respiratory changes in response to cognitive load: A systematic review. *Neural Plasticity*, *2016,* Article 8146809. https://doi.org/10.1155/2016/8146809

Grassmann, M., Vlemincx, E., von Leupoldt, A., & 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. *Applied Ergonomics*, *59*, 274–282. https://doi.org/10.1016/j.apergo.2016.09.006

Grueschow, M., Kleim, B., & Ruff, C. C. (2020). Role of the locus coeruleus arousal system in cognitive control. *Journal of Neuroendocrinology*, *32*(12), Article 12890. https://doi.org/10.1111/jne.12890

Gwizdka, J. (2021). “Overloading” Cognitive (Work)Load: What Are We Really Measuring? In F. D. Davis, R. Riedl, J. vom Brocke, P. M. Léger, A. B. Randolph, G. Müller-Putz (Eds.) *Information Systems and Neuroscience. NeuroIS 2021. Lecture Notes in Information Systems and Organisation, Vol. 52*. Springer, Cham. https://doi.org/10.1007/978-3-030-88900-5_9

Hagemann, D., & Naumann, E. (2001). The effects of ocular artifacts on (lateralized) broadband power in the EEG. *Clinical Neurophysiology*, *112*(2), 215–231. https://doi.org/10.1016/S1388-2457(00)00541-1

Haith, A. M., & Krakauer, J. W. (2018). The multiple effects of practice: skill, habit and reduced cognitive load. *Current Opinion in Behavioral Sciences*, *20*, 196–201. https://doi.org/10.1016/j.cobeha.2018.01.015

Hammel, K. R., Fisher, D. L., & 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*, *46*(26), 2174–2178.

Hancock, P. A., & Matthews, G. (2019). Workload and Performance: Associations, Insensitivities, and Dissociations. *Human Factors*, *61*(3), 374–392. https://doi.org/10.1177/0018720818809590

Hancock, P. A., & Sawyer, B. D. (2015). Judging Thieves of Attention: Commentary on "Assessing Cognitive Distraction in the Automobile, " by Strayer, Turrill, Cooper, Coleman, Medeiros-Ward, and Biondi (2015). *Human Factors*, *57*(8), 1339–1342. https://doi.org/10.1177/0018720815578971

Harbluk, J. L., Noy, Y. I., Trbovich, P. L. & Eizenman, M. (2007). An on-road assessment of cognitive distraction: Impacts on drivers’ visual behavior and braking performance. *Accident Analysis & Prevention, 39*(2), 372–379. https://doi.org/10.1016/j.aap.2006.08.013

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, *52*, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Hecht, T., Darlagiannis, E., & Bengler, K. (2020). Non-driving Related Activities in Automated Driving – An Online Survey Investigating User Needs. In T. Ahram, W. Karwowski, S. Pickl, & R. Taiar (Eds.), *Human Systems Engineering and Design II. IHSED 2019. Advances in Intelligent Systems and Computing* (Vol. 1026, pp. 182–188). Springer, Cham. https://doi.org/10.1007/978-3-030-27928-8_28

Heine, T., Lenis, G., Reichensperger, P., Beran, T., Doessel, O., & Deml, B. (2017). Electrocardiographic features for the measurement of drivers' mental workload. *Applied Ergonomics*, *61*, 31–43. https://doi.org/10.1016/j.apergo.2016.12.015

Helfrich, R. F., & Knight, R. T. (2016). Oscillatory Dynamics of Prefrontal Cognitive Control. *Trends in Cognitive Sciences*, *20*(12), 916–930. https://doi.org/10.1016/j.tics.2016.09.007

Hertzum, M. (2021). Associations among workload dimensions, performance, and situational characteristics: a meta-analytic review of the Task Load Index. *Behaviour and Information Technology*. https://doi.org/10.1080/0144929X.2021.2000642

Hickman, J. S., Hanowski, R. J., & Bocanegra, J. (2010). *Distraction in commercial trucks and buses: Assessing prevalence and risk in conjunction with crashes and near-crashes* (Report No. FMCSA-RRR-10-049). U.S. Department of Transportation, Federal Motor Carrier Safety Administration. https://doi.org/10.21949/1502693

Highway Loss Data Institute. (2020). Compendium of HLDI collision avoidance research. *HLDI Bulletin, 37*(12).

Hollander, J., & Huette, S. (2022). Extracting blinks from continuous eye-tracking data in a mind wandering paradigm. *Consciousness and Cognition*, *100,* Article 103303. https://doi.org/10.1016/j.concog.2022.103303

Hopstaken, J. F., van der Linden, D., Bakker, A. B., & Kompier, M. A. J. (2015). The window of my eyes: Task disengagement and mental fatigue covary with pupil dynamics. *Biological Psychology*, *110*, 100–106. https://doi.org/10.1016/j.biopsycho.2015.06.013

Hsu, C.-F., Propp, L., Panetta, L., Martin, S., Dentakos, S., Toplak, M. E., & Eastwood, J. D. (2018). Mental effort and discomfort: Testing the peak-end effect during a cognitively demanding task. *PLoS ONE*, *13*(2), Article e0191479. https://doi.org/10.1371/journal.pone.0191479

Hu, X., & 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. *Journal of Safety Research*, *72*, 173–187. https://doi.org/https://doi.org/10.1016/j.jsr.2019.12.015

Hughes, A. M., Hancock, G. M., Marlow, S. L., Stowers, K., & Salas, E. (2019). Cardiac Measures of Cognitive Workload: A Meta-Analysis. *Human Factors*, *61*(3), 393–414. https://doi.org/10.1177/0018720819830553

International Organization for Standardization. (2016). Road vehicles - Transport information and control systems - Detection-response task (DRT) for assessing attentional effects of cognitive load in driving (ISO 17488:2016).

International Organization for Standardization. (2010). Road vehicles - Ergonomic aspects of transport information and control systems - Simulated lane change test to assess in-vehicle secondary task demand (ISO 26022:2010)

Islam, M. K., Rastegarnia, A., & Yang, Z. (2016). Methods for artifact detection and removal from scalp EEG: A review. *Clinical Neurophysiology*, *46*(4-5), 287–305. https://doi.org/10.1016/j.neucli.2016.07.002

Ismail, L. E., & Karwowski, W. (2020). Applications of EEG indices for the quantification of human cognitive performance: A systematic review and bibliometric analysis. *PLoS ONE*, *15*(12), Article e0242857. https://doi.org/10.1371/journal.pone.0242857

Itoh, K., Miki, Y., Kubo, N., Takeda, Y., & Tanaka, H. (2006). A study on estimating the variation of driver's state by EEGs and EOGs. *SAE Technical Papers* (2006-01-0575). https://doi.org/10.4271/2006-01-0575

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*(4), 394–412. https://doi.org/10.1080/09658211003702171

Jamshed, S. (2014). Qualitative research method-interviewing and observation. *Journal of Basic and Clinical Pharmacy*, *5*(4), 87–88.

Jarosch, O., Paradies, S., Feiner, D., & Bengler, K. (2019). Effects of non-driving related tasks in prolonged conditional automated driving – A Wizard of Oz on-road approach in real traffic environment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *65*, 292–305. https://doi.org/10.1016/j.trf.2019.07.023

Johnson, M. J., Chahal, T., Stinchcombe, A., Mullen, N., Weaver, B., & Bédard, M. (2011). Physiological responses to simulated and on-road driving. *International Journal of Psychophysiology*, *81*(3), 203–208. https://doi.org/10.1016/j.ijpsycho.2011.06.012

Joos, K., Gilles, A., Van de Heyning, P., De Ridder, D., & Vanneste, S. (2014). From sensation to percept: The neural signature of auditory event-related potentials. *Neuroscience and Biobehavioral Reviews*, *42*, 148–156. https://doi.org/10.1016/j.neubiorev.2014.02.009

Jorlöv, S., Bohman, K., & Larsson, A. (2017). Seating positions and activities in highly automated cars - A qualitative study of future automated driving scenarios. *Conference Proceedings International Research Council on the Biomechanics of Injury, IRCOBI*, 13–22.

Joshi, S., & Gold, J. I. (2020). Pupil Size as a Window on Neural Substrates of Cognition. *Trends in Cognitive Sciences*, *24*(6), 466–480. https://doi.org/10.1016/j.tics.2020.03.005

Kato, Y., Endo, H., & Kizuka, T. (2009). Mental fatigue and impaired response processes: Event-related brain potentials in a Go/NoGo task. *International Journal of Psychophysiology*, *72*(2), 204–211. https://doi.org/10.1016/j.ijpsycho.2008.12.008

Keren, G. (2014). Between- or Within-Subjects design: A Methodological Dilemma. In G. Keren & C. Lewis (Eds.), *A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues* (pp. 257–272).

Khan, M. Q., & Lee, S. (2019). Gaze and eye tracking: Techniques and applications in ADAS. *Sensors*, *19*(24), Article 5540. https://doi.org/10.3390/s19245540

Kihlstrom, J. F. (2021). Ecological Validity and "Ecological Validity." *Perspectives on Psychological Science*, *16*(2), 466–471. https://doi.org/10.1177/1745691620966791

Kircher, K., & Ahlstrom, C. (2017). Minimum Required Attention: A Human-Centered Approach to Driver Inattention. *Human Factors*, *59*(3), 471–484. https://doi.org/10.1177/0018720816672756

Klauer, S. G., Klauer, S. G., Dingus, T. a., Dingus, T. a., Neale, V. L., Neale, V. L., Sudweeks, J. D., Sudweeks, J. D., Ramsey, D. J., & Ramsey, D. J. (2006). *The Impact of Driver Inattention On Near Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data* (Report No. DOT HS 810 594). U.S. Department of Transportation, National Highway Traffic Safety Administration.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews*, *29*(2-3), 169–195. https://doi.org/10.1016/S0165-0173(98)00056-3

Kluwe-Schiavon, B., Viola, T. W., Sanvicente-Vieira, B., Malloy-Diniz, L. F., & Grassi-Oliveira, R. (2017). Balancing automatic-controlled behaviors and emotional-salience states: A dynamic executive functioning hypothesis. *Frontiers in Psychology*, *7*, Article 2067. https://doi.org/10.3389/fpsyg.2016.02067

Knapen, T., De Gee, J. W., Brascamp, J., Nuiten, S., Hoppenbrouwers, S., & Theeuwes, J. (2016). Cognitive and ocular factors jointly determine pupil responses under equiluminance. *PLoS ONE*, *11*(5), Article e0155574. https://doi.org/10.1371/journal.pone.0155574

Köhler, A.-L., Pelzer, J., Seidel, K., & Ladwig, S. (2019). Sitting Postures and Activities in Autonomous Vehicles – New Requirements towards Occupant Safety. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 63*(1), 1874–1878. https://doi.org/10.1177/1071181319631327

König, M., & Neumayr, L. (2017). Users' resistance towards radical innovations: The case of the self-driving car. *Transportation Research Part F: Traffic Psychology and Behaviour*, *44*, 42–52. https://doi.org/10.1016/j.trf.2016.10.013

Kountouriotis, G. K., & Merat, N. (2016). Leading to distraction: Driver distraction, lead car, and road environment. *Accident Analysis and Prevention*, *89*, 22–30. https://doi.org/10.1016/j.aap.2015.12.027

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

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, *84*(3), 394–421. https://doi.org/10.1016/j.biopsycho.2010.03.010

Kurzban, R. (2010). Does the brain consume additional glucose during self-control tasks? *Evolutionary Psychology*, *8*(2), 244–259. https://doi.org/10.1177/147470491000800208

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, *25*(3), 296–324. https://doi.org/10.1016/j.preteyeres.2006.01.002

Langer, I., Abendroth, B., & Bruder, R. (2016). Driver condition detection. In H. Winner, C. Singer, S. Hakuli, & F. Lotz (Eds.), *Handbook of Driver Assistance Systems* (pp. 871–889). Springer Cham. https://doi.org/10.1007/978-3-319-12352-3

Large, D. R., Burnett, G., Salanitri, D., Lawson, A., & Box, E. (2019). A longitudinal simulator study to explore drivers' behaviour in level 3 automated vehicles. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI,* 222–232. https://doi.org/10.1145/3342197.3344519

Lee, J. D. (2014). Dynamics of Driver Distraction: The process of engaging and disengaging. *Annals of Advances in Automotive Medicine*, *58*, 24–32.

Lee, J. D., Caven, B., Haake, S., & Brown, T. L. (2001). Speech-based interactions with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. *Human Factors*, *43*(4), 631–640. https://doi.org/10.1518/001872001775870340

Lee, J. D., Roberts, S. C., Hoffman, J. D., & Angell, L. S. (2012). Scrolling and driving: How an MP3 player and its aftermarket controller affect driving performance and visual behavior. *Human Factors*, *54*(2), 250–263. https://doi.org/10.1177/0018720811429562

Lehtonen, E., Lappi, O., & Summala, H. (2012). Anticipatory eye movements when approaching a curve on a rural road depend on working memory load. *Transportation Research Part F: Traffic Psychology and Behaviour*, *15*(3), 369–377. https://doi.org/10.1016/j.trf.2011.08.007

Leisman, G., Moustafa, A. A., & Shafir, T. (2016). Thinking, Walking, Talking: Integratory Motor and Cognitive Brain Function. *Frontiers in Public Health*, *4*, Article 94. https://doi.org/10.3389/fpubh.2016.00094

Leledakis, A., Östh, J., Davidsson, J., & Jakobsson, L. (2021). The influence of car passengers' sitting postures in intersection crashes. *Accident Analysis and Prevention*, *157,* Article 106170. https://doi.org/10.1016/j.aap.2021.106170

Leonhardt, S., Leicht, L., & Teichmann, D. (2018). Unobtrusive vital sign monitoring in automotive environments—A review. *Sensors*, *18*(9), Article 3080. https://doi.org/10.3390/s18093080

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences*, *6*(7), 299–305. https://doi.org/10.1016/S1364-6613(02)01923-X

Liang, Y., & Lee, J. D. (2010). Combining cognitive and visual distraction: Less than the sum of its parts. *Accident Analysis and Prevention*, *42*(3), 881–890. https://doi.org/10.1016/j.aap.2009.05.001

Lin, J. Y., Franconeri, S., & Enns, J. T. (2008). Objects on a collision path with the observer demand attention. *Psychological Science*, *19*(7), 686–692. https://doi.org/10.1111/j.1467-9280.2008.02143.x

Lindgren, T., Pink, S., & Fors, V. (2021). Fore-sighting autonomous driving - An Ethnographic approach. *Technological Forecasting and Social Change*, *173,* Article 121105. https://doi.org/10.1016/j.techfore.2021.121105

Lobjois, R., Faure, V., Désiré, L., & 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? *Safety Science*, *134,* Article 105046. https://doi.org/10.1016/j.ssci.2020.105046

Loeches De La Fuente, H., Berthelon, C., Fort, A., Etienne, V., De Weser, M., Ambeck, J., & Jallais, C. (2019). Electrophysiological and performance variations following driving events involving an increase in mental workload. *European Transport Research Review*, *11,* Article 42. https://doi.org/10.1186/s12544-019-0379-z

Lohani, M., Payne, B. R., & Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Frontiers in Human Neuroscience*, *13,* Article 57. https://doi.org/10.3389/fnhum.2019.00057

Longo, L. (2018). On the reliability, validity and sensitivity of three mental workload assessment techniques for the evaluation of instructional designs: A case study in a third-level course. *Proceedings of the 10th International Conference on Computer Supported Education*, *1,* 166–178. https://doi.org/10.5220/0006801801660178

Luong, T., Martin, N., Argelaguet, F., & Lecuyer, A. (2019). Studying the mental effort in virtual versus real environments. *2019 IEEE Conference on Virtual Reality and 3D User Interfaces*, 809–816. https://doi.org/10.1109/VR.2019.8798029

Magliacano, A., Fiorenza, S., Estraneo, A., & Trojano, L. (2020). Eye blink rate increases as a function of cognitive load during an auditory oddball paradigm. *Neuroscience Letters*, *736,* Article 135293. https://doi.org/10.1016/j.neulet.2020.135293

Malta, L., Ljung Aust, M., Faber, F., Metz, B., Saint Pierre, G., Benmimoun, M., & Schäfer, R. (2012). *EuroFOT Deliverable 6.4 Final results: Impacts on traffic safety*.

Mantzke, O., & Keinath, A. (2015). Relating the Detection Response Task to Critical Events – Consequences of high Cognitive Workload to Brake Reaction Times. *Procedia Manufacturing*, *3*, 2381–2386. https://doi.org/10.1016/j.promfg.2015.07.386

Markkula, G., Engström, J., Lodin, J., Bärgman, J., & Victor, T. (2016). A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies. *Accident Analysis and Prevention*, *95*, 209–226. https://doi.org/10.1016/j.aap.2016.07.007

Marti, P., Jallais, C., Koustanaï, A., Guillaume, A., & Mars, F. (2022). Impact of the driver's visual engagement on situation awareness and takeover quality. *Transportation Research Part F: Traffic Psychology and Behaviour*, *87*, 391–402. https://doi.org/10.1016/j.trf.2022.04.018

Mathissen, M., Hennes, N., Faller, F., Leonhardt, S., & Teichmann, D. (2021). Investigation of Three Potential Stress Inducement Tasks During On-Road Driving. *IEEE Transactions on Intelligent Transportation Systems*, *23*(5), 4823–4832. https://doi.org/10.1109/TITS.2021.3112811

Matthews, G., & Campbell, S. E. (2009). Sustained performance under overload: Personality and individual differences in stress and coping. *Theoretical Issues in Ergonomics Science*, *10*(5), 417–422. https://doi.org/10.1080/14639220903106395

Matthews, G., Reinerman-Jones, L. E., Barber, D. J., & Abich, J. (2015). The psychometrics of mental workload: Multiple measures are sensitive but divergent. *Human Factors*, *57*(1), 125–143. https://doi.org/10.1177/0018720814539505

Matthews, G., Wohleber, R., Lin, J., Funke, G., & Neubauer, C. (2019). Monitoring task fatigue in contemporary and future vehicles: A review. In D. Cassenti (Eds.) *Advances in Human Factors in Simulation and Modeling. AHFE 2018. Advances in Intelligent Systems and Computing*, *780*, 101–112. https://doi.org/10.1007/978-3-319-94223-0_10

Medeiros-Ward, N., Cooper, J. M., & Strayer, D. L. (2014). Hierarchical Control and driving. *Journal of Experimental Psychology: General*, *143*(3), 953–958. https://doi.org/10.1037/a0035097

Mehler, B., Reimer, B., & 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. *Human Factors*, *54*(3), 396–412. https://doi.org/10.1177/0018720812442086

Mehler, B., Reimer, B., Coughlin, J. F., & Dusek, J. A. (2009). Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. *Transportation Research Record: Journal of the Transportation Research Board*, *2138*(1), 6–12. https://doi.org/10.3141/2138-02

Mehler, B., Reimer, B., & Dusek, J. A. (2011). *MIT AgeLab Delayed Digit Recall Task (n-back)* (Working Paper 2011-3B).

Mehler, B., Reimer, B., & Zec, M. (2012). Defining workload in the context of driver state detection and HMI evaluation. *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 187–191. https://doi.org/10.1145/2390256.2390288

Merat, N., & Jamson, A. H. (2008). The effect of stimulus modality on signal detection: Implications for assessing the safety of in-vehicle technology. *Human Factors*, *50*(1), 145–158. https://doi.org/10.1518/001872008X250656

Merat, N., Kountouriotis, G., Tomlinson, A., Carsten, O., & Engström, J. (2015). Performance on the Detection Response Task during driving: Separating the manual and cognitive element of the secondary task. *Proceedings of the 4th International Driver Distraction and Inattention Conference*.

Meyer, D. E., & Kieras, D. E. (1997). A Computational Theory of Executive Cognitive Processes and Multiple-Task Performance: Part 1. Basic Mechanisms. *Psychological Review, 104*(1), 3-65. https://doi.org/10.1037/0033-295X.104.1.3

Miller, D., Sun, A., Johns, M., Ive, H., Sirkin, D., Aich, S., & Ju, W. (2015). Distraction becomes engagement in automated driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 59(1), 1676–1680. https://doi.org/10.1177/1541931215591362

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Milleville-Pennel, I., & Charron, C. (2015). Driving for Real or on a Fixed-Base Simulator: Is It so Different? An Explorative Study. *Presence*, *24*(1), 74–91. https://doi.org/10.1162/PRES_a_00216

Mitchell, D. J., McNaughton, N., Flanagan, D., & Kirk, I. J. (2008). Frontal-midline theta from the perspective of hippocampal "theta." *Progress in Neurobiology*, *86*(3), 156–185. https://doi.org/10.1016/j.pneurobio.2008.09.005

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Morando, A., Victor, T., & Dozza, M. (2020). A Bayesian Reference Model for Visual Time-Sharing Behaviour in Manual and Automated Naturalistic Driving. *IEEE Transactions on Intelligent Transportation Systems*, *21*(2), 803–814. https://doi.org/10.1109/TITS.2019.2900436

Morando, M. M., Tian, Q., Truong, L. T., & Vu, H. L. (2018). Studying the Safety Impact of Autonomous Vehicles Using Simulation-Based Surrogate Safety Measures. *Journal of Advanced Transportation*, *2018,* Article 6135183. https://doi.org/10.1155/2018/6135183

Mueller, J. A. (2015). Driving in a Simulator Versus on-Road: the Effect of Increased Mental Effort While Driving on Real Roads and a Driving Simulator. [Doctoral dissertation, Montana state university]. https://scholarworks.montana.edu/xmlui/handle/1/9068

Mulder, L. J. M. (1992). Measurement and analysis methods of heart rate and respiration for use in applied environments. *Biological Psychology*, *34*(2–3), 205–236. https://doi.org/10.1016/0301-0511(92)90016-N

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences*, *15*(10), 453–459. https://doi.org/10.1016/j.tics.2011.07.011

Murphey, Y., Kochhar, D. S., & Xie, Y. (2019). Driver workload in an autonomous vehicle. *SAE Technical Papers*, *2019-01–0872*. https://doi.org/10.4271/2019-01-0872

Muttart, J. W., Fisher, D. L., Knodler, M., & Pollatsek, A. (2007). Driving without a clue: Evaluation of driver simulator performance during hands-free cell phone operation in a work zone. *Transportation Research Record: Journal of the Transportation Research Board, 2018*(1), 9-14. https://doi.org/10.3141/2018-02

Nakano, T., Kato, M., Morito, Y., Itoi, S., & Kitazawa, S. (2013). Blink-related momentary activation of the default mode network while viewing videos. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(2), 702–706. https://doi.org/10.1073/pnas.1214804110

Náñez, J. (1988). Perception of Impending Collision in 3-to 6-Week-Old Human Infants. *Infant Behavior and Development*, *11*(4), 447–463. https://doi.org/10.1016/0163-6383(88)90005-7

Niezgoda, M., Tarnowski, A., Kruszewski, M., & 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. *Transportation Research Part F: Traffic Psychology and Behaviour*, *32*, 23–34. https://doi.org/10.1016/j.trf.2015.04.012

Nilsson, E., Ahlström, C., Barua, S., Fors, C., Lindén, P., Svanberg, B., Begum, S., Uddin Ahmed, M., & Anund, A. (2017). *Vehicle Driver Monitoring – Sleepiness and Cognitive load* (VTI report 937A).

Niu, J., & Ma, C. (2022). Is it Good or Bad to Provide Driver Fatigue Warning During Take-Over in Highly Automated Driving? *Transportation Research Record: Journal of the Transportation Research Board*, *2676*(2), 762–774. https://doi.org/10.1177/03611981211046920

Noh, Y., Kim, S., Jang, Y. J., & Yoon, Y. (2021). Modeling individual differences in driver workload inference using physiological data. *International Journal of Automotive Technology*, *22*, 201–212. https://doi.org/10.1007/s12239-021-0020-8

Nunes, L., & Recarte, M. A. (2002). Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *5*(2), 133–144. https://doi.org/10.1016/S1369-8478(02)00012-8

O'Donnell, C. R. D., & Eggemeier, T. F. (1986). Workload assessment methodology. In *Handbook of Perception and Human Performance* (2nd ed.).

Oei, N. Y. L., Veer, I. M., Wolf, O. T., Spinhoven, P., Rombouts, S. A. R. B., & Elzinga, B. M. (2012). Stress shifts brain activation towards ventral "affective" areas during emotional distraction. *Social Cognitive and Affective Neuroscience*, *7*(4), 403–412. https://doi.org/10.1093/scan/nsr024

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). *Driver Distraction in Commercial Vehicle Operations* (Report No. FMCSA-RRR-09-042). U.S. Department of Transportation, Federal Motor Carrier Safety Administration. https://doi.org/10.21949/1502647

Pace-Schott, E. F., Amole, M. C., Aue, T., Balconi, M., Bylsma, L. M., Critchley, H., Demaree, H. A., Friedman, B. H., Gooding, A. E. K., Gosseries, O., Jovanovic, T., Kirby, L. A. J., Kozlowska, K., Laureys, S., Lowe, L., Magee, K., Marin, M. F., Merner, A. R., Robinson, J. L., … VanElzakker, M. B. (2019). Physiological feelings. *Neuroscience and Biobehavioral Reviews*, *103*, 267–304. https://doi.org/10.1016/j.neubiorev.2019.05.002

Patel, A. N., Howard, M. D., Roach, S. M., Jones, A. P., Bryant, N. B., Robinson, C. S. H., Clark, V. P., & Pilly, P. K. (2018). Mental State Assessment and Validation Using Personalized Physiological Biometrics. *Frontiers in Human Neuroscience*, *12,* Article 221. https://doi.org/10.3389/fnhum.2018.00221

Patten, C. J. D., Kircher, A., Östlund, J., & Nilsson, L. (2004). Using mobile telephones: Cognitive workload and attention resource allocation. *Accident Analysis and Prevention*, *36*(3), 341–350. https://doi.org/10.1016/S0001-4575(03)00014-9

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, 35, 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525

Pfleging, B., Rang, M., & Broy, N. (2016). Investigating user needs for non-driving-related activities during automated driving. *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia*, 91–99. https://doi.org/10.1145/3012709.3012735

Pfurtscheller, G., Stancák A. Jr., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band - an electrophysiological correlate of cortical idling: a review. *International Journal of Psychophysiology, 24*(1-2), 39–46. https://doi.org/10.1016/S0167-8760(96)00066-9

Putze, F., Jarvis, J. P., & Schultz, T. (2010). Multimodal recognition of cognitive workload for multitasking in the car. *Proceedings of the 20th International Conference on Pattern Recognition*, 3748–3751. https://doi.org/10.1109/ICPR.2010.913

Qu, S. Q., & Dumay, J. (2011). The qualitative research interview. *Qualitative Research in Accounting and Management*, *8*(3), 238–264. https://doi.org/10.1108/11766091111162070

Quintana, D. S., & Heathers, J. A. J. (2014). Considerations in the assessment of heart rate variability in biobehavioral research. *Frontiers in Psychology*, *5,* Article 805. https://doi.org/10.3389/fpsyg.2014.00805

Rahman, H., Ahmed, M. U., Barua, S., & Begum, S. (2020). Non-contact-based driver's cognitive load classification using physiological and vehicular parameters. *Biomedical Signal Processing and Control*, *55*, Article 101634. https://doi.org/10.1016/j.bspc.2019.101634

Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology: Applied*, *6*(1), 31–43. https://doi.org/10.1037/1076-898X.6.1.31

Recarte, M. A., & Nunes, L. M. (2003). Mental Workload While Driving: Effects on Visual Search, Discrimination, and Decision Making. *Journal of Experimental Psychology: Applied*, *9*(2), 119–137. https://doi.org/10.1037/1076-898X.9.2.119

Recarte, M. A., Pérez, E., Conchillo, Á., & Nunes, L. M. (2008). Mental Workload and Visual Impairment: Differences between Pupil, Blink, and Subjective Rating. *The Spanish Journal of Psychology*, *11*(2), 374–385. https://doi.org/10.1017/s1138741600004406

Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis and Prevention*, *43*(5), 1771–1781. https://doi.org/10.1016/j.aap.2011.04.008

Reimer, B. (2009). Impact of cognitive task complexity on drivers' visual tunneling. *Transportation Research Record*, *2138*(1), 13–19. https://doi.org/10.3141/2138-03

Reimer, B., & Mehler, B. (2011). The impact of cognitive workload on physiological arousal in young adult drivers: A field study and simulation validation. *Ergonomics*, *54*(10), 932–942. https://doi.org/10.1080/00140139.2011.604431

Reimer, B., Mehler, B., Dobres, J., & Coughlin, J. F. (2013). The Effects of a Production Level "Voice-Command" Interface on Driver Behavior: Reported Workload, Physiology, Visual Attention, and Driving Performance (Technical Report 2013-17A). MIT AgeLab.

Reimer, B., Mehler, B., Wang, Y., & 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. *Human Factors*, *54*(3), 454–468. https://doi.org/10.1177/0018720812437274

Rensink, R. A. (2009). Attention: Change Blindness and Inattentional Blindness. In W. P. Banks (Ed.), *Encyclopedia of Consciousness* (pp. 47–59). Academic Press. https://doi.org/10.1016/B978-012373873-8.00006-2

Richter, M., & Slade, K. (2017). Interpretation of physiological indicators of motivation: Caveats and recommendations. *International Journal of Psychophysiology*, *119*, 4–10. https://doi.org/10.1016/j.ijpsycho.2017.04.007

Ronen, A., & Yair, N. (2013). The adaptation period to a driving simulator. *Transportation Research Part F: Traffic Psychology and Behaviour*, *18*, 94–106. https://doi.org/10.1016/j.trf.2012.12.007

Rosey, F., & Auberlet, J. M. (2014). Driving simulator configuration impacts drivers' behavior and control performance: An example with studies of a rural intersection. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 99–111. https://doi.org/10.1016/j.trf.2014.09.009

Ross, J. A., & Van Bockstaele, E. J. (2021). The Locus Coeruleus- Norepinephrine System in Stress and Arousal: Unraveling Historical, Current, and Future Perspectives. *Frontiers in Psychiatry*, *11*, Article 601519. https://doi.org/10.3389/fpsyt.2020.601519

Rowley, J. (2012). Conducting research interviews. *Management Research Review*, *35*(3–4), 260–271. https://doi.org/10.1108/01409171211210154

Rukonic, L., Pungu Mwange, M.-A., & Kieffer, S. (2021). UX Design and Evaluation of Warning Alerts for Semi-autonomous Cars with Elderly Drivers. *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 25-36. https://doi.org/10.5220/0010237000250036

Sadaghiani, S., & Kleinschmidt, A. (2016). Brain Networks and α-Oscillations: Structural and Functional Foundations of Cognitive Control. *Trends in Cognitive Sciences*, *20*(11), 805–817. https://doi.org/10.1016/j.tics.2016.09.004

Salvucci, D. D., & Beltowska, J. (2008). Effects of memory rehearsal on driver performance: Experiment and theoretical account. *Human Factors*, *50*(5), 834–844. https://doi.org/10.1518/001872008X354200

Sato, T., Takeda, Y., Akamatsu, M., & Kitazaki, S. (2020). Evaluation of driver drowsiness while using automated driving systems on driving simulator, test course and public roads. In H. Krömker (Ed.), *HCI in Mobility, Transport, and Automotive Systems. Driving Behavior, Urban and Smart Mobility. HCII 2020. Lecture Notes in Computer Science, vol 12213,* Springer, Cham. https://doi.org/10.1007/978-3-030-50537-0_7

Sauder, K. A., Johnston, E. R., Skulas-Ray, A. C., Campbell, T. S., & West, S. G. (2012). Effect of meal content on heart rate variability and cardiovascular reactivity to mental stress. *Psychophysiology*, *49*(4), 470–477. https://doi.org/10.1111/j.1469-8986.2011.01335.x

Sawyer, B. D., Lee, J., Dobres, J., Mehler, B., Coughlin, J. F., & Reimer, B. (2016). Effects of a Voice interface on mirror check decrements in older and younger multitasking drivers. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 60*(1), 16–20. https://doi.org/10.1177/1541931213601004

Scarapicchia, V., Brown, C., Mayo, C., & Gawryluk, J. R. (2017). Functional magnetic resonance imaging and functional near-infrared spectroscopy: Insights from combined recording studies. *Frontiers in Human Neuroscience*, *11,* Article 419. https://doi.org/10.3389/fnhum.2017.00419

Schmidt, J., Laarousi, R., Stolzmann, W., & Karrer-Gauß, K. (2018). Eye blink detection for different driver states in conditionally automated driving and manual driving using EOG and a driver camera. *Behavior Research Methods*, *50*, 1088–1101. https://doi.org/10.3758/s13428-017-0928-0

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66. https://doi.org/10.1037/0033-295X.84.1.1

Schömig, N., Hargutt, V., Neukum, A., Petermann-Stock, I., & Othersen, I. (2015). The Interaction Between Highly Automated Driving and the Development of Drowsiness. *Procedia Manufacturing*, *3*, 6652–6659. https://doi.org/10.1016/j.promfg.2015.11.005

Schoofs, D., Preuß, D., & Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology*, *33*(5), 643–653. https://doi.org/10.1016/j.psyneuen.2008.02.004

Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed method research design. *Kolner Zeitschrift Fur Soziologie Und Sozialpsychologie*, *69*, 107–131. https://doi.org/10.1007/s11577-017-0454-1

Schwabe, L. (2017). Memory under stress: from single systems to network changes. *European Journal of Neuroscience*, *45*(4), 478–489. https://doi.org/10.1111/ejn.13478

Sescousse, G., Ligneul, R., van Holst, R. J., Janssen, L. K., de Boer, F., Janssen, M., Berry, A. S., Jagust, W. J., & Cools, R. (2018). Spontaneous eye blink rate and dopamine synthesis capacity: preliminary evidence for an absence of positive correlation. *European Journal of Neuroscience*, *47*(9), 1081–1086. https://doi.org/10.1111/ejn.13895

Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, *5*, Article 258. https://doi.org/10.3389/fpubh.2017.00258

Shariff, A., Bonnefon, J. F., & Rahwan, I. (2017). Psychological roadblocks to the adoption of self-driving vehicles. *Nature Human Behaviour*, *1*, 694–696. https://doi.org/10.1038/s41562-017-0202-6

Shiffrin, R. M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review*, *84*(2), 127-190. https://doi.org/10.1037/0033-295X.84.2.127

Shinar, D. (2015). Cognitive Workload ≠ Crash Risk: Rejoinder to Study by Strayer et al. (2015). *Human Factors*, *57*(8), 1328–1330. https://doi.org/10.1177/0018720815574943

Shinar, D. (2017). Research methods. In D. Shinar (Ed.), *Traffic safety and human behavior* (2nd ed., pp. 41–86). Emerald Publisher.

Shine, J. M., Bissett, P. G., Bell, P. T., Koyejo, O., Balsters, J. H., Gorgolewski, K. J., Moodie, C. A., & Poldrack, R. A. (2016). The Dynamics of Functional Brain Networks: Integrated Network

States during Cognitive Task Performance. *Neuron*, *92*(2), 544–554. https://doi.org/10.1016/j.neuron.2016.09.018

Shulman, R. G., Rothman, D. L., Behar, K. L., & Hyder, F. (2004). Energetic basis of brain activity: Implications for neuroimaging. *Trends in Neurosciences*, *27*(8), 489–495. https://doi.org/10.1016/j.tins.2004.06.005

Sidikova, M., Martinek, R., Kawala-Sterniuk, A., Ladrova, M., Jaros, R., Danys, L., & Simonik, P. (2020). Vital sign monitoring in car seats based on electrocardiography, ballistocardiography and seismocardiography: A review. *Sensors*, *20*(19), Article 5699. https://doi.org/10.3390/s20195699

Singh, S. (2015). *Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey* (Report No. DOT HS 812 115). U.S. Department of Transportation, National Highway Traffic Safety Administration. Traffic Safety Facts - Crash Stats. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812506

Solovey, E. T., Zec, M., Perez, E. A. G., Reimer, B., & Mehler, B. (2014). Classifying driver workload using physiological and driving performance data: Two field studies. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 4057–4066. https://doi.org/10.1145/2556288.2557068

SAE International. (2021). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (J2016_202104). https://www.sae.org/standards/content/j3016_202104/

Stemmler, G., & Wacker, J. (2010). Personality, emotion, and individual differences in physiological responses. *Biological Psychology*, *84*(3), 541–551. https://doi.org/10.1016/j.biopsycho.2009.09.012

Strayer, D. L., Cooper, J. M., & Drews, F. A. (2004). What do Drivers Fail to See When Conversing on a Cell Phone? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 48*(19), 2213–2217. https://doi.org/10.1177/154193120404801902

Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J., Medeiros-Ward, N., & Biondi, F. (2013). *Measuring cognitive distraction in the automobile* (Technical Report). AAA Foundation for Traffic Safety.

Strayer, D. L., & Fisher, D. L. (2016). SPIDER: A Framework for Understanding Driver Distraction. *Human Factors*, *58*(1), 5–12. https://doi.org/10.1177/0018720815619074

Strayer, D. L., Turrill, J., Cooper, J. M., Coleman, J. R., Medeiros-Ward, N., & Biondi, F. (2015). Assessing Cognitive Distraction in the Automobile. *Human Factors*, *57*(8), 1300–1324. https://doi.org/10.1177/0018720815575149

Streff, F., & Spradlin, H. (2000). *Driver distraction, aggression, and fatigue: a synthesis of the literature and guidelines for Michigan planning* (Technical Report No. UMTRI-2000-10). Michigan Office of Highway Safety Planning.

Sullivan, J. M., Flannagan, M. J., Pradhan, A. K., & Bao, S. (2016). *Literature Review of Behavioral Adaptations to Advanced Driver Assistance Systems* (Technical Report). AAA Foundation for Traffic Safety.

Summala, H. (2007). Towards Understanding Motivational and Emotional Factors in Driver Behaviour: Comfort Through Satisficing. In P. C. Cacciabue (Ed.), *Modelling Driver Behaviour in Automotive Environments* (pp. 189–207). Springer. https://doi.org/10.1007/978-1-84628-618-6

Szalma, J. L. (2009). Individual differences in performance, workload, and stress in sustained attention: Optimism and pessimism. *Personality and Individual Differences*, *47*(5), 444–451. https://doi.org/10.1016/j.paid.2009.04.019

Takeda, Y., Yoshitsugu, N., Itoh, K., & Kanamori, N. (2012). Assessment of Attentional Workload while Driving by Eye-fixation-related Potentials. *Kansei Engineering International Journal*, *11*(3), 121–126. https://doi.org/10.5057/kei.11.121

Tao, D., Tan, H., Wang, H., Zhang, X., Qu, X., & Zhang, T. (2019). A systematic review of physiological measures of mental workload. *International Journal of Environmental Research and Public Health*, *16*(15), Article 2716. https://doi.org/10.3390/ijerph16152716

Tatum, IV, W. O. (2014). Handbook of EEG interpretation. In W. O. Tatum, IV (Ed.), *Handbook of EEG Interpretation* (2nd ed.). Springer Publishing Company.

Tennant, C., Stares, S., & Howard, S. (2019). Public discomfort at the prospect of autonomous vehicles: Building on previous surveys to measure attitudes in 11 countries. *Transportation Research Part F: Traffic Psychology and Behaviour*, *64*, 98–118. https://doi.org/10.1016/j.trf.2019.04.017

Tivesten, E., & Dozza, M. (2014). Driving context and visual-manual phone tasks influence glance behavior in naturalistic driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *26*, 258–272. https://doi.org/10.1016/j.trf.2014.08.004

Tivesten, E., & Dozza, M. (2015). Driving context influences drivers' decision to engage in visual-manual phone tasks: Evidence from a naturalistic driving study. *Journal of Safety Research*, *53*, 87–96. https://doi.org/10.1016/j.jsr.2015.03.010

Tivesten, E., Morando, A., & Victor, T. (2015). The timecourse of driver visual attention in naturalistic driving with adaptive cruise control and forward collision warning. *4th International Driver Distraction and Inattention Conference*, 1–12.

Todd, J. J., Fougnie, D., & Marois, R. (2005). Visual short-term memory load suppresses temporo-parietal junction activity and induces inattentional blindness. *Psychological Science*, *16*(12), 965–972. https://doi.org/10.1111/j.1467-9280.2005.01645.x

Tomasi, D., Chang, L., Caparelli, E. C., & Ernst, T. (2007). Different activation patterns for working memory load and visual attention load. *Brain Research*, *1132*, 158–165. https://doi.org/10.1016/j.brainres.2006.11.030

Törnros, J. E. B., & Bolling, A. K. (2005). Mobile phone use - Effects of handheld and handsfree phones on driving performance. *Accident Analysis and Prevention*, *37*(5), 902–909. https://doi.org/10.1016/j.aap.2005.04.007

Tortora, G. J., & Derrickson, B. (2007). *Principles of anatomy and physiology* (11th ed.). John Wiley & Sons, Inc.

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology*, *8*, Article 1454. https://doi.org/10.3389/fpsyg.2017.01454

Umetani, K., Singer, D. H., McCraty, R., & Atkinson, M. (1998). Twenty-four hour time domain heart rate variability and heart rate: Relations to age and gender over nine decades. *Journal of the American College of Cardiology*, *31*(3), 593–601. https://doi.org/10.1016/S0735-1097(97)00554-8

United Nations Economic Commission for Europe. (2020). Uniform provisions concerning the approval of vehicles with regard to Automated Lane Keeping Systems (ECE/TRANS/WP.29/2020/81).

United Nations. (2015). Sustainable Development Goals. Goal 3: Ensure healthy lives and promote well-being for all at all ages. https://www.un.org/sustainabledevelopment/health/

Van Acker, B. B., Parmentier, D. D., Vlerick, P., & Saldien, J. (2018). Understanding mental workload: from a clarifying concept analysis toward an implementable framework. *Cognition, Technology and Work*, *20*, 351–365. https://doi.org/10.1007/s10111-018-0481-3

Van der Heiden, R. M. A., Kenemans, J. L., Donker, S. F., & Janssen, C. P. (2021). The Effect of Cognitive Load on Auditory Susceptibility During Automated Driving. *Human Factors*. https://doi.org/10.1177/0018720821998850

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin and Review*, *25*, 2005–2015. https://doi.org/10.3758/s13423-018-1432-y

Vetrugno, R., Liguori, R., Cortelli, P., & Montagna, P. (2003). Sympathetic skin response: Basic mechanisms and clinical applications. *Clinical Autonomic Research*, *13*, 256–270. https://doi.org/10.1007/s10286-003-0107-5

Victor, T., Dozza, M., Bärgman, J., Boda, C.-N., Engström, J., Flannagan, C., Lee, J. D., & Markkula, G. (2015). *Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk, SHRP2 Safety Research* (Report S2-S08A-RW-1). https://doi.org/10.17226/22297

Victor, T., Harbluk, J., & Engström, J. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*(2), 167–190. https://doi.org/10.1016/j.trf.2005.04.014

Victor, T., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human Factors*, *60*(8), 1095–1116. https://doi.org/10.1177/0018720818788164

Vogelpohl, T., Kühn, M., Hummel, T., & Vollrath, M. (2019). Asleep at the automated wheel—Sleepiness and fatigue during highly automated driving. *Accident Analysis and Prevention*, *126*, 70–84. https://doi.org/10.1016/j.aap.2018.03.013

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

Vossel, S., Geng, J. J., & Fink, G. R. (2014). Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles. *The Neuroscientist*, *20*(2), 150–159. https://doi.org/10.1177/1073858413494269

Wandtner, B., Schömig, N., & Schmidt, G. (2018). Secondary task engagement and disengagement in the context of highly automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *58*, 253–263. https://doi.org/10.1016/j.trf.2018.06.001

Wang, C. A., & Munoz, D. P. (2015). A circuit for pupil orienting responses: Implications for cognitive modulation of pupil size. *Current Opinion in Neurobiology*, *33*, 134–140. https://doi.org/10.1016/j.conb.2015.03.018

Wang, J., Warnecke, J. M., Haghi, M., & Deserno, T. M. (2020). Unobtrusive health monitoring in private spaces: The smart vehicle. *Sensors*, *20*(9), Article 2442. https://doi.org/10.3390/s20092442

Wang, Y., Reimer, B., Dobres, J., & Mehler, B. (2014). The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand. *Transportation Research Part F: Traffic Psychology and Behaviour*, *26*, 227–237. https://doi.org/10.1016/j.trf.2014.08.003

Watson, A. B., & Yellott, J. I. (2012). A unified formula for light-adapted pupil size. *Journal of Vision*, *12*(10), 1–16. https://doi.org/10.1167/12.10.12

Weaver, B. W., & DeLucia, P. R. (2020). A Systematic Review and Meta-Analysis of Takeover Performance During Conditionally Automated Driving. *Human Factors*. https://doi.org/10.1177/0018720820976476

Westbrook, A., & Braver, T. S. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective and Behavioral Neuroscience*, *15*, 395–415. https://doi.org/10.3758/s13415-015-0334-y

Wever, M., Smeets, P., & Sternheim, L. (2015). Neural correlates of intolerance of uncertainty in clinical disorders. *Journal of Neuropsychiatry and Clinical Neurosciences*, *27*(4), 345–353. https://doi.org/10.1176/appi.neuropsych.14120387

Wheatley, C. L., Esplin, J., Loveless, S. M., Cooper, J. M., Biondi, F., & Strayer, D. L. (2018). Performance and workload trends: The effects of repeated exposure to "high" demand tasks. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 62(1), 6–10. https://doi.org/10.1177/1541931218621002

Wiberg, H., Nilsson, E., Lindén, P., Svanberg, B., & Poom, L. (2015). Physiological responses related to moderate mental load during car driving in field conditions. *Biological Psychology*, *108*, 115–125. https://doi.org/10.1016/j.biopsycho.2015.03.017

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*(3), 449–455. https://doi.org/10.1518/001872008X288394

Wickens, J. R., Horvitz, J. C., Costa, R. M., & Killcross, S. (2007). Dopaminergic mechanisms in actions and habits. *Journal of Neuroscience*, *27*(31), 8181–8183. https://doi.org/10.1523/JNEUROSCI.1671-07.2007

Wijayaratna, K. P., Cunningham, M. L., Regan, M. A., Jian, S., Chand, S., & Dixit, V. V. (2019). Mobile phone conversation distraction: Understanding differences in impact between simulator and naturalistic driving studies. *Accident Analysis and Prevention*, *129*, 108–118. https://doi.org/10.1016/j.aap.2019.04.017

Winkeler, M., Erander, H., Buvat, J., Sengupta, A., Mehl, R., Sule, S., KVJ, S., & Khemka, Y. (2019). *The Autonomous Car: A Consumer Perspective*. Capgemini Research Institute. https://www.capgemini.com/insights/research-library/steering-the-future-of-the-autonomous-car/

Winn, B., Whitaker, D., Elliott, D. B., & Phillips, N. J. (1994). Factors affecting light-adapted pupil size in normal human subjects. *Investigative Ophthalmology and Visual Science*, *35*(3), 1132–1137.

Wintersberger, P., Schartmüller, C., & Shadeghian-borojeni, S. (2021). Evaluation of Imminent Take-Over Requests With Real Automation on a Test Track. *Human Factors*. https://doi.org/10.1177/00187208211051435

Wood, G., Hartley, G., Furley, P. A., & Wilson, M. R. (2016). Working Memory Capacity, Visual Attention and Hazard Perception in Driving. *Journal of Applied Research in Memory and Cognition*, *5*(4), 454–462. https://doi.org/10.1016/j.jarmac.2016.04.009

Wood, M., Robbel, P., Maas, M., Tebbens, R. D., Meijs, M., Harb, M., Reach, J., Robinson, K., Wittmann, D., Srivastava, T., Essayed Bouzouraa, M., Liu, S., Wang, Y., Knobel, C., Boymanns, D., Löhning, M., Delhink, B., Kaule D., Krüger, R., … Schlicht, P. (2019). *Safety first for automated driving*. https://www.connectedautomateddriving.eu/wp-content/uploads/2019/09/Safety_First_for_Automated_Driving.pdf

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, *72*(8), 2031–2046. https://doi.org/10.3758/APP.72.8.2031

World Health Organization. (2018). *Global status report on road safety*. https://www.who.int/publications/i/item/9789241565684

Wusk, G., & Gabler, H. (2018). Non-invasive detection of respiration and heart rate with a vehicle seat sensor. *Sensors*, *18*(5), Article 1463. https://doi.org/10.3390/s18051463

Wynne, R. A., Beanland, V., & Salmon, P. M. (2019). Systematic review of driving simulator validation studies. *Safety Science*, *117*, 138–151. https://doi.org/10.1016/j.ssci.2019.04.004

Young, K. L., Lenné, M. G., & Williamson, A. R. (2011). Sensitivity of the lane change test as a measure of in-vehicle system demand. *Applied Ergonomics*, *42*(4), 611–618. https://doi.org/10.1016/j.apergo.2010.06.020

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: mental workload in ergonomics. *Ergonomics*, *58*(1), 1–17. https://doi.org/10.1080/00140139.2014.956151

Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. *Human Factors*, *44*(3), 365–375. https://doi.org/10.1518/0018720024497709

Zheng, X. S., McConkie, G. W., & Simons, D. J. (2005). Effects of verbal and spatial-imagery tasks on drivers' situation awareness of traffic information. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 49*(22), 1999–2003. https://doi.org/10.1177/154193120504902224

Zijlstra, F. R. H. (1993). *Efficiency in Work Behavior: A Design Approach for Modern Tools* [Doctoral dissertation, Technische Universiteit Delft, the Netherlands]. http://resolver.tudelft.nl/uuid:d97a028b-c3dc-4930-b2ab-a7877993a17f