# Applied and laboratory-based autonomic and neurophysiological monitoring during sustained attention tasks

By Alice Stephenson

A thesis submitted in partial fulfilment of the requirements of the University of the West of England, Bristol for the degree of Doctor of Philosophy

This research programme was carried out in collaboration with Airbus Group Innovations

Faculty of Health and Applied Sciences

University of the West of England

January 2021

# Abstract

Fluctuations during sustained attention can cause momentary lapses in performance which can have a significant impact on safety and wellbeing. However, it is less clear how unrelated tasks impact current task processes, and whether potential disturbances can be detected by autonomic and central nervous system measures in naturalistic settings. In a series of five experiments, I sought to investigate how prior attentional load impacts semi-naturalistic tasks of sustained attention, and whether neurophysiological and psychophysiological monitoring of continuous task processes and performance could capture attentional lapses. The **first experiment** explored various non-invasive electrophysiological and subjective methods during multitasking. The **second experiment** employed a manipulation of multitasking, task switching, to attempt to unravel the negative lasting impacts of multitasking on neural oscillatory activity, while the **third experiment** employed a similar paradigm in a semi-naturalistic environment of simulated driving. The **fourth experiment** explored the feasibility of measuring changes in autonomic processing during a naturalistic sustained monitoring task, autonomous driving, while the **fifth experiment** investigated the visual demands and acceptability of a biological based monitoring system. The results revealed several findings. While the **first experiment** demonstrated that only self-report ratings were able to successfully disentangle attentional load during multitasking; the **second and third experiment** revealed deficits in parieto-occipital alpha activity and continuous performance depending on the attentional load of a previous unrelated task. The **fourth experiment** demonstrated increased sympathetic activity and a smaller distribution of fixations during an unexpected event in autonomous driving, while the **fifth experiment** revealed the acceptability of a biological based monitoring system although further research is needed to unpick the effects on attention. Overall, the results of this thesis help to provide insight into how autonomic and central processes manifest during semi-naturalistic sustained attention tasks. It also provides support for a neuro- or biofeedback system to improve safety and wellbeing.

# Acknowledgements

First and foremost, I would like to acknowledge and thank my Director of Studies, Dr Chris Alford, for all his help and assistance with this thesis. I would also like to thank the rest of my supervisory team Dr James Macdonald, Dr Kris Kinsey, and Dr Adrian Kendrick, for their insightful conversations, MATLAB support, and continual feedback. My thanks also extend to Jan, Hardeep, and Hannah, who dealt with my never-ending need for equipment and lab space.

I want to take this opportunity to express my gratitude towards the Faculty of Environment and Technology and the Bristol Robotics Laboratory, especially: Professor Praminda Caleb-Solley, Dr Iveta Eimontaite, Professor Tony Pipe, the technicians, the assisted living team, and the human-robot interaction team. By welcoming me with open arms into your department and laboratory provided me with a scarce opportunity to develop as a researcher and be part of insightful multidisciplinary research beyond these pages.

# Collaborative Work

Chapter 5 and Chapter 6 stemmed from the FLOURISH project. FLOURISH was developed in response to Innovate UK's Connected and Autonomous Vehicles Collaboration Research and Development competition. A three-year project adopted a user-centred approach to better understand consumer demands and expectations for autonomous vehicle technology in older adults. The FLOURISH project enabled the inclusion of the simulator as well as the autonomous vehicle used in this PhD research. The opportunity for the PhD developed from FLOURISH and was co-funded by partners Airbus Group Innovations and UWE Research Business and Innovation.

For Chapter 5, the autonomous Pod and technical support was provided by the Transport Systems Catapult, developed by RDM/Aurrigo. The human-machine interface was designed and developed by Designability and integrated with the Pod by Transport Systems Catapult and Dave Barnett (Jolly Good Engineering Ltd.). The larger UWE and Cardiff research team including Dr Iveta Eimontaite, Dr Chris Alford, Professor Praminda Caleb-Solly, and Professor Phillip Morgan, were involved in the study design and implementation of the study. I was co-responsible for data collection, led by Dr Iveta Eimontaite, and supported by research placement students Tabasum Khatun and Joseph Davis. For this Chapter, I was responsible for determining the research questions, data analysis, including all pre-processing, statistical analysis, the production of all figures and tables, and interpretation of the findings. Dr Iveta Eimontaite, Dr Chris Alford, and Professor Praminda Caleb-Solly aided in the interpretation of the findings. All writings are my own. The work in this Chapter has been published in *Frontiers in Psychology,* reference provided on page 5.

For Chapter 6, the human-machine interface was designed and developed by Designability and integrated with the simulator by Bristol Robotics Laboratory and Dave Barnett (Jolly Good Engineering Ltd.). The simulated driving scenario was developed and designed in combination with Bristol Robotics Laboratory and Transport Systems Catapult. I was responsible for determining the research questions, running the study, collecting the data, analysing the data, including all pre-processing, statistical analysis, and the production of all figures and tables, and interpretation of the findings. All writings are my own.

# Publications and Presentations

**Journal articles:**

Stephenson, A., Eimontaite, I., Caleb-Solly, P., Morgan, P., L., Khatun, T., Davis, J., & Alford, C. (2020). Effects of an unexpected and expected event on older adults' autonomic arousal and eye fixations during autonomous driving. *Frontiers in Psychology*, *11,* 57196

**Conference papers:**

Stephenson, A., Eimontaite, I., Caleb-Solly, P., & Alford, C. (2020). The impact of a biological driver state monitoring system on visual attention during partially automated driving. In: Stanton N. (eds) *Advances in Human Aspects of Transportation*. Advances in Intelligent Systems and Computing, vol 1212. Springer, Cham

**Presentations:**

Stephenson, A., Eimontaite, I., Caleb-Solly, P., & Alford, C. (Dec 2019). *Measuring cognitive enhancement through psychophysiological measures during autonomous driving.* Oral presentation for an invited talk at Swinburne University. Melbourne, Australia

Stephenson A., Macdonald, J., Kinsey, K., & Alford, C. (Nov 2019). *The effect of prior perceptual load on visuomotor neural oscillatory mechanisms and task performance.* Oral presentation at Australasian Cognitive Neuroscience Association Meeting. Tasmania, Australia

Stephenson A., Macdonald, J., Kinsey, K., & Alford, C. (Sep 2019). *The effect of task load on neural oscillatory activity when switching between a visuospatial and continuous visuomotor task.* Oral presentation at BPS Psychobiology Annual Meeting. Windermere, UK

Stephenson A., Macdonald, J., Kinsey, K., & Alford, C. (Apr 2019). *The effect of task load on the temporal EEG dynamics of switching between a visuospatial task and a continuous visuomotor task.* Poster presentation at British Neuroscience Association Festival of Neuroscience. Dublin, Ireland

Stephenson A., Kendrick, A., Kinsey, K., & Alford, C. (Sep 2018). *The impact of acute stress on neural and psychophysiological mechanisms during a cognitive stress task.* Poster presentation at BPS Psychobiology Annual Meeting. Windermere, UK

# Contents

# List of Tables

# List of Figures

# List of Appendices

# 1.0 General Introduction

Sustained attention is a fundamental construct associated with the capacity to respond and maintain performance over prolonged periods of time (Robertson & O'Connell, 2012). Impairments in the ability to select relevant stimuli and ignore irrelevant stimuli occur in everyday situations such as driving, working, and studying, and impacts general cognitive abilities such as memory and learning (e.g. Awh et al., 2006; Chun, 2011). Consequently, decrements in sustained attention pose a serious threat to safety and wellbeing in naturalistic settings. For instance, fatigue and drowsiness have both been shown to impact sustained attention, suppressing driver ability and control of the vehicle (e.g. Gunzelmann et al., 2011), and situation awareness (e.g. Smith et al., 2009), resulting in critical, life-threatening situations such as road accidents (Connor et al., 2002). As such, it has been argued that it is necessary to develop practical and safe interventions to improve sustained attention ability. As models of attention have demonstrated a complex and overlapping system that fluctuate over periods of time (e.g. Esterman & Rothlein, 2019), continual monitoring of physiological indices with a high temporal resolution may reveal the mechanisms that underlie attentional lapses, as well as provide preventative measures by assessing when they are more likely to occur. Monitoring of autonomic and central nervous system mechanisms may therefore provide the potential to improve sustained attention ability, having a wide range of implications for improving safety and wellbeing in everyday situations and clinical populations.

This Chapter will first begin with introducing key theoretical and neuroanatomical models of selective and sustained attention. Given the fluctuating nature of sustained attention, several physiological indices that have demonstrated their ability to capture deficits will be subsequently introduced. The Chapter will then discuss attentional deficits during human-machine interactions and clinical populations, which may benefit from autonomic and neurophysiological monitoring. In view of the significant implications for physiological monitoring during sustained attention tasks for safety and wellbeing in naturalistic settings, several issues concerning laboratory-based studies are discussed, followed by a brief thesis overview.

## 1.1 Attention

Attention is considered a function that refers to a set of mechanisms related to orienting, focusing, and selecting (Nobre & Mesulam, 2014). Humans successfully identify and select goal-relevant information against competing distractors with limited cognitive resources in complex dynamic environments (Lavie & Dalton, 2014). Selective attention serves as the mechanism which allows us to resolve this competition by selectively filtering task-relevant information from task-irrelevant information (Posner, 1980). Selective attention can be deployed in two ways. Orienting attention can happen overtly, which enables eyes and head movements towards the target stimulus; or covertly, attending to the target when there is not enough time to execute eye movements. Another important aspect is that reorienting can happen quickly and effortlessly, known as exogenous; or slowly and deliberately, known as endogenous (Posner, 1980). While selective attention is often measured in a transient manner, sustained attention captures how performance fluctuates over time, when tasks can be effortful and demanding, or monotonous and undemanding (Robertson & O'Connell, 2012). This first section of the Chapter begins with introducing the early versus late selection models of selective attention, followed by underload and overload models attempting to capture sustained attention depicting performance fluctuations over time. Arousal and vigilance mechanisms are then introduced as fundamental elements of sustained attention. Finally, the section reviews the evidence from sophisticated neuroimaging techniques that have advanced understanding and informed theoretical models of attention, revealing dynamic interactions involving inhibition and excitation across a distributed network during sustained attention tasks.

### 1.1.1 Selective attention

Attentional research has attempted to uncover the mechanisms underpinning selective attention. The locus of selection was first proposed by Broadbent (1958) as an early selection process. Broadbent's bottleneck model suggests that numerous sensory inputs are filtered for further processing based on their physical characteristics. This filter prevents the information-processing system from becoming overloaded. As the unprocessed information is rejected at an early stage, these are removed from any further processing. According to early selection models, attention occurs at an early perceptual level before stimulus identification. Soon after, Moray (1959) demonstrated that participants perceived information in the unattended channel when preceded by the participant's name. Consequently, Moray (1959) argued that unattended stimuli can be processed semantically. As such, irrelevant information is attenuated, rather than eliminated (Treisman, 1964). These findings drove the development of the late selection view: Deutsch & Deutsch (1963) argued that the locus of selection is a late process as perception has an unlimited capacity and attention operates on the result of stimulus identification. Filtering of irrelevant information happens only after all inputs have been recognised.

While the early versus late selection models were the primary debates of the mid-1900s, a dominant theory to surface in the 1990s suggested that selective attention is dependant on the processing demands of the current task, which determines the efficiency of distractor rejection (first proposed in Lavie & Tsal, 1994). The perceptual load theory has been hugely influential and proposes a solution to the early versus late selection debate and relates to limited capacity models of information processing. When a task imposes high perceptual load, capacity is reached and distractors cannot be processed, resulting in performance that is consistent with early selection. However, when a task involves low perceptual load, all available stimuli are processed (i.e. distractors and targets), necessitating late selection. If perceptual load is high, cognitive load impacts the allocation of attention, and if there is a high working memory requirement, this can cause late selection to fail (Lavie, 2005). Therefore, both types of selection can occur, depending on the perceptual properties of the task. In this case, perceptual load considers the perceptual properties of the task only (e.g. colour of the stimuli) to aid information processing, whereas cognitive load includes both bottom-up and top-down processing required to process relevant information, including working memory (Macdonald & Lavie, 2011). Although it is now mostly agreed that the locus of selection is not exclusively early or late (e.g. Allport et al., 1994), many have criticised the hybrid perceptual load model. Some researchers have suggested that rather than any perceptual or cognitive load impact, the heterogeneous stimuli in a high load condition 'dilutes' any effect of distractor items (Benoni & Tsal, 2010); under low load, the distractor is generally presented alone, while under high load, the distractor is presented with other distractors. Therefore, when more stimuli are competing for processing, there is less available capacity (Benoni & Tsal, 2010; Tsal & Benoni, 2010a, 2010b; Wilson et al., 2011). Others have also argued that it is not directly perceptual load elements that limit attentional processing. The biased competition account of attentional selection (Desimone & Duncan, 1995) argues that all objects are competing for processing. This competition can be biased by bottom-up mechanisms such as the visual salience of the stimuli, or top-down control mechanisms such as motivational behaviour. For example, spatial proximity between target and distractor has been shown to impact distractor interference. Paquet & Craig (1997) found that selective attention was negatively impacted under low load for near but not far distractors. In addition, Linnell & Caparos (2011) found that high perceptual load reduced distractor interference when cognitive load was low. Evidently, a variety of factors in combination with perceptual and cognitive load can influence selective attention.

### 1.1.2 Sustained attention

Robertson & O'Connell (2012) proposed that sustained attention 'is the capacity to maintain accurate responding over time across tasks which can be effortful and demanding, or monotonous and undemanding'. As a result, sustained attention can be considered an add on to selective attention: relevant stimuli continue to be selected and irrelevant stimuli continue to be ignored (Robertson & O'Connell, 2012). A critical element of sustained attention is its temporal limits, i.e. how engagement is maintained over an extended period of time. Despite some researchers putting time limits on sustained attention (e.g. Langner & Eickhoff (2013) meta-analysis only included studies of tasks lasting at least 10 seconds), there is no agreement that task engagement evolves into sustained attention after a set duration. As such, this thesis defines sustained attention as a non-specific construct associated with the capacity to maintain attention and provide continuous effort on task demands. Although sustained attention is less studied then transient aspects of attention, performance measures have found that decrements in sustained attention fluctuate but generally increase over time and are often characterised by increased reaction times and/or increased errors. In particular, the mean error rate (e.g. Robertson et al., 1997) and the trial-to-trial variability in reaction times (e.g. Esterman et al., 2013) have been used to capture fluctuations over a task period.

Lapses in sustained attention are most commonly associated with either cognitive overload or cognitive underload accounts. The overload account suggests that attentional resources decline over time (Hitchcock et al., 1999). Sustaining attention is cognitively demanding due to the continuous processing of stimuli and response selection. Therefore, attentional lapses reflect limits in available cognitive resources (Head & Helton, 2014). Underload theories suggest that task demands are under stimulating and lead to task disengagement (Nachreiner & Hänecke, 1992). Research has demonstrated that monotonous, unchallenging tasks can lead to reduced performance and are perceived as highly demanding (e.g. Robertson & O'Connell, 2012; Warm et al., 2008). This has been associated with mind wandering specifically, suggesting that individuals experience a state in which cognitive processing is driven by internally oriented goals, rather than responding to task-relevant demands (Cheyne et al., 2009). Yet, inconsistent findings in the literature cast doubt over both models as the overload resource depletion model cannot fully account for decreasing task demands that lead to reduced performance (e.g. Ralph et al., 2017), and the underload model cannot account for increasing task demands leading to reduced performance (e.g. Puma et al., 2018).

A novel and integrative model attempts to explain the discrepancies between the overload and underload models of sustained attention. The resource-control theory (Thomson et al., 2015) suggests that there is a bias for attentional resources towards mind wandering as this is the default state of individuals. Therefore, attentional resources do not decrease over time, but more attentional resources are devoted to mind wandering. In some cases, mind wandering will not result in performance costs if the task does not require all attentional resources. Therefore, attention can successfully be divided between internal and task-related thoughts. The model also highlights the importance of executive control, preventing resources being unequally allocated towards task-irrelevant thoughts, yet this reduces over time. Therefore, mind wandering consumes resources that can cause failures in executive control, and as executive control decreases over time, this will result in increased feelings of greater effort. This model arguably explains the literature better than the limited overload and underload accounts and has been supported by studies employing reward paradigms, demonstrating that reward manipulations modulate overall performance. For example, Esterman et al. (2016) found that a monetary reward improved performance of a continuous performance task. The model postulates that task engagement can alleviate performance decrements as it enables efficient resource distribution. More difficult tasks result in greater performance decrements as more attentional resources must be devoted to the task, and so reductions in executive control will have a larger impact over time. However, these findings also suggest rewards improve performance during sustained attention, implying that a depletion of cognitive resources can be reversed. The opportunity cost model could explain this finding (Kurzban et al., 2013). The model posits that the stronger the perceived benefits of alternate tasks relative to the current task, the greater the perceived effort. In summary, performance fluctuations during sustained attention tasks have been associated with motivation and reward processing, rather than simply an allocation of cognitive resources.

An integrated model of sustained attention must also consider the resources of arousal and vigilance as sustained attention shares a significant relationship with both. These concepts are often used interchangeably, and although they do overlap, there are clear differences between the three. Arousal is considered a global energetic state of the system, whereas sustained attention is considered a state to be maintained (Sarter et al., 2001). Arousal is a cognitive resource rather than a cognitive function and can be considered a baseline amount of resources available for a task (Esterman & Rothlein, 2019). Like sustained attention, arousal shares a relationship with performance over time. Hebb's (1955) work based upon the Yerkes-Dodson Law (Yerkes & Dodson, 1908) described the relationship between arousal and performance as an inverted U-shaped curve. Performance is best at an intermediate level, and both under- and over-arousal states are associated with poor performance. A number of circuits and corresponding neurotransmitters are associated with an arousal network (e.g. Szabadi, 2015), though low tonic locus coeruleus noradrenergic system has been associated with low task engagement due to hypoarousal, and high tonic locus coeruleus activity has been associated with low task engagement due to hyperarousal (Aston-Jones & Cohen, 2005). These fluctuations in locus coeruleus noradrenergic neuromodulation impact frontoparietal regions which are thought to modulate sustained attention (Grefkes et al., 2010; Lenartowicz et al., 2013; further discussed in Chapter 1.1.3). Like arousal, vigilance is also considered a global energetic state of the system. As a component of sustained attention, vigilance is described as a readiness to perceive and respond to detect infrequent and unpredictable targets or events (e.g. Mackworth, 1948; Robertson & O'Connell, 2012). So although vigilance decrements over time are characterised similarly to sustained attention by increased reaction times and/or errors, they are defined as the decreased probability of detecting target stimuli or events as time-on-task increases (e.g. Mackworth, 1948). As such, vigilance represents a global mechanism with a slower temporal resolution to sustained attention. Therefore, vigilance decrements can be distinguished from sustained attention decrements by their distinct temporal characteristics: where sustained attention can fluctuate over seconds, vigilance decays are apparent only after 30 minutes under waking conditions (Robertson & O'Connell, 2012).

### 1.1.3 Neuroanatomical basis

Advancements in functional neuroimaging in the 1990s provided the opportunity to improve understanding of the temporal dynamics of attentional selection. Understanding the neural networks and dynamic interactions that support sustained attention may provide insight into the mechanisms that underlie attentional lapses, which could inform new theoretical models of sustained attention deficits and provide frameworks for behavioural effects. While a dominant theory described the structural and functional basis of individual features of attention (Petersen & Posner, 2012; Posner & Petersen, 1990), further work suggested that sustained attention is reliant on a diverse network of cortical structures and dynamic interactions between distinct systems (Fortenbaugh et al., 2017). These models are discussed below.

An influential theory of the neuroanatomical and functional basis of attention was formulated by Posner (1980) and later expanded on owing to the advent on neuroimaging and statistical methods (Petersen & Posner, 2012; Posner & Petersen, 1990). The authors discussed three functionally and structurally distinct systems: an alerting system, an orienting system, and an executive system. These three systems have been associated with both selective and sustained attention. Whereas selective attention may include both the orienting and executive system, sustained attention may include alerting and orienting (Tang et al., 2015), as well as the executive network (Langner & Eickhoff, 2013).

The alerting system concerns the state of arousal in maintaining attention to facilitate accurate responses. In an experimental situation, a phasic shift in alertness is most typically measured by presenting a warning cue before the onset of a stimulus (Petersen & Posner, 2012). The cue changes the state by enabling preparation for the response to a stimulus. As such, when a cue precedes a target, response times are reduced. Alertness is modulated by a range of elements, such as the time of day (Posner, 1975), length of the task (Posner & Petersen, 1990), as well as the general level of wakefulness (Szabadi, 2015). The alerting network comprises right thalamic, frontal and parietal regions, and is influenced by the cortical noradrenergic system that arises from the locus coeruleus (Aston-Jones & Cohen, 2005; Petersen & Posner, 2012; Posner & Petersen, 1990). The evidence suggests that tonic alertness is heavily lateralised to the right hemisphere (Sturm & Willmes, 2001); whereas phasic alertness has been associated with left hemisphere mechanisms (Ivry & Robertson, 1998).

The second system, the orienting system, refers to the process of selecting the relevant sensory stimulus. There are three critical operations to this response. Firstly, the individual must disengage from the attended location. Secondly, they must shift their attention to a new location. Finally, they must engage towards the new location. The most widely utilised task for assessing reorientation is the Posner cueing task (Posner, 1980). A cue provides spatial information about the upcoming target. The cue can be either valid or invalid. If invalid, the target appears in a different location to the cue, and so attention is reoriented towards the location of the target. Reaction times increase following an invalid cue (Posner, 1980). A large body of research has demonstrated that the reorienting subcomponents are associated with distinct brain structures and have separate effects on attention, activating two separated but internally correlated networks: the dorsal attention network (DAN), and the ventral attention network (VAN; Corbetta et al., 2008; Petersen & Posner, 2012). The DAN is related to a goal-directed selection of sensory stimuli during focused attention, whereas the VAN detects salient and behaviourally relevant stimuli triggering a shift in attention. In addition to visual areas, the dorsal system comprises of bilateral frontoparietal regions the intraparietal sulcus, superior parietal lobule, and frontal eye fields. Once a shift of attention occurs, activation of the ventral stream including anatomical areas the temporoparietal junction and the ventral frontal cortex ensues. It has been thought that the ventral system is lateralised to the right hemisphere (Corbetta et al., 2008), however recent research has disputed this claim including several studies that have demonstrated bilateral temporoparietal junction activation during reorienting (e.g. Doricchi et al., 2010; Vossel et al., 2009). Moreover, the ventral stream has been associated with evaluative processes, rather than the processes that precede target detection (Doricchi et al., 2010; Han & Marois, 2014). Nonetheless, an interlinked network of frontoparietal areas seem to be associated with attention reorienting.

Finally, the executive network refers to the processes following stimulus detection which encompasses two main systems (Dosenbach et al., 2008). The first system concerns the sustained state, whereas the second system concerns a phasic state to adapt to current task demands. For maintaining a state over the task period, the cingulo-opercular control network consisting of the dorsal anterior cingulate cortex, the adjacent medial superior frontal cortex, and bilateral anterior insula, is activated (Dosenbach et al., 2006). A frontoparietal system, in contrast, is implemented during phasic adjustments, activating lateral frontal and parietal regions (Dosenbach et al., 2006, 2007), though these are distinct to frontoparietal orienting networks (Petersen & Posner, 2012). Therefore, the executive network is considered a stable system that can adapt to sudden changes supporting sustained performance (Dosenbach et al., 2008).

Further research has attempted to delineate temporal and structural activity by identifying large-scale distributed networks involved in sustained attention: the default mode network and a frontoparietal control network. The default mode network consists of numerous structures that are activated during passive rest and mind wandering, including regions along the anterior and posterior midline, the medial temporal lobe, and the lateral parietal cortex (Buckner et al., 2008). Although initially supposed that the default mode network was task-irrelevant (e.g. Fox et al., 2005), it is now thought to play a complex role supporting cognition (e.g. Lin et al., 2017; for a review, see Raichle, 2015). A suppression of the network is commonly described during task engagement (e.g. Greicius & Menon, 2004; Singh & Fawcett, 2008). Research has also indicated a significant role of the frontoparietal control network, including areas of the DAN and VAN (Fortenbaugh et al., 2017). In agreement with the role of the frontoparietal attention network and the default mode network during sustained attention, Langner & Eickhoff's (2013) meta-analysis identified several cortical and subcortical structures consistently activated during vigilant attention tasks, including areas of the default mode network, the DAN and VAN, as well as the executive control attention networks identified by Petersen & Posner (2012). The role of the frontoparietal and default mode network is also supported by models of sustained attention. Resource control theory introduced earlier (Chapter 1.1.2) suggests that the frontoparietal control network, reflecting attention, and the default mode network, reflecting mind wandering, should be inversely related and fluctuate over time representing sustained attention (Thomson et al., 2015). Esterman et al., (2013) revealed that performance (lower response variability) was associated with increased blood-oxygen-level-dependent (BOLD) signal in the default mode network, whereas BOLD was less pronounced in the DAN. When participants performed worse, BOLD activity increased in the DAN, yet was less pronounced in the default mode network. Although these results seem slightly contradictory, as the attention network is important for supporting attentional processes, this has led to the default mode interference hypothesis suggesting that disorders of attention, mind wandering, and performance deficits, are associated with failures to effectively suppress activity within the core default mode regions (Broyd et al., 2009). Overall, these results suggest that both the frontoparietal and default mode network have a role in integrating attentional processes during sustained attention.

Taken together, the literature demonstrates that sustained attention relies on a large-scale network of cortical areas, particularly a frontoparietal network responsible for the representation of relevant information, and the default mode network responsible for suppression of irrelevant information. As multiple elements including properties of external stimuli and internal states play a role in mediating performance over time, there are extensive complex overlaps between the circuits that mediate the different elements of attention. Furthermore, it is now well recognised that performance fluctuations are multifaceted and oscillate over multiple timescales from seconds (e.g. phasic alertness) to days (e.g. circadian) to years (e.g. homeostasis), and so the temporal dynamics of global states of arousal and vigilance are also fundamental components of sustained attention.

## 1.2 Autonomic and neurophysiological monitoring

The above-mentioned studies describe how fluctuations in sustained attention are associated with a complex array of elements including cognitive resource allocation, motivation, reward processing, arousal, and vigilance, which all operate on a series of timescales. Sustained attention is not a unitary phenomenon and consequently no single measure can inform us about the independent attentional mechanisms that operate within different subsystems. Considering that lapses in attention fluctuate over time, continual monitoring of physiological indices with good temporal resolution may reveal the mechanisms that underlie attentional lapses, as well as provide preventative measures by assessing when they are more likely to occur. This section will introduce several reliable physiological and neurophysiological measures that have been employed to measure sustained attention in the literature. These measures will be explored further in subsequent Chapters.

### 1.2.1 Electrocardiography (ECG)

The autonomic nervous system (ANS) plays a vital role in the modulation of physiological arousal to meet current environmental demands (Thayer & Lane, 2000). The ANS is comprised of the sympathetic nervous system (SNS), a quick-acting system associated with energy mobilisation; and the parasympathetic nervous system (PNS), a slower response system associated with energy inhibition. Both branches work together to maintain homeostasis. An inability of the ANS to respond appropriately to changes in environmental demands may leave humans more prone to distractions (Thayer & Lane, 2000). As such, fluctuations in sustained attention may represent dysregulation of the ANS. In addition, given the high sensitivity of the ANS to arousal level, the ANS is an attractive target for the development of a real-time measure of attentional monitoring.

The cardiovascular system is under control of both the SNS and PNS (Gordan et al., 2015). Three major components, the P-wave, QRS complex and the T-wave, form the cardiac cycle and are associated with different phases of contraction of the heart (see Figure 1.1). The R-wave is the most dominant wave of the ECG and is produced by depolarisation of the main mass of the ventricles. Heart rate is the total number of R-waves over a one-minute period (i.e. beats per minute). An increase in heart rate can be a result of either an increase in SNS activity, a decrease in PNS activity, or a combination of both (Thayer et al., 2010). Another popular approach for analysing cardiovascular activity is heart rate variability (HRV): the variation in the time interval between consecutive R-waves in milliseconds. There are over 70 variables of HRV and analyses can be performed in the time-domain, the frequency-domain, and with non-linear indices (Laborde et al., 2017). It is most common to analyse in the time-domain, calculating features based upon the time between two successive R-waves, known as the RR interval. For example, the root-mean-square of successive differences between RR intervals (RMSSD), the standard deviation of all normal RR intervals (SDNN), and percentage of successive normal RR intervals more than 50ms (pNN50). While RMSSD and pNN50 reflect vagal tone (PNS), SDNN reflects all the cyclic components responsible for variability in the recording. ECG data can also be transformed into the frequency-domain and non-linear indices, however, some of these parameters have come under fire recently as it is not clear what some represent and standards are lacking, particularly for non-linear indices (Laborde et al., 2017).

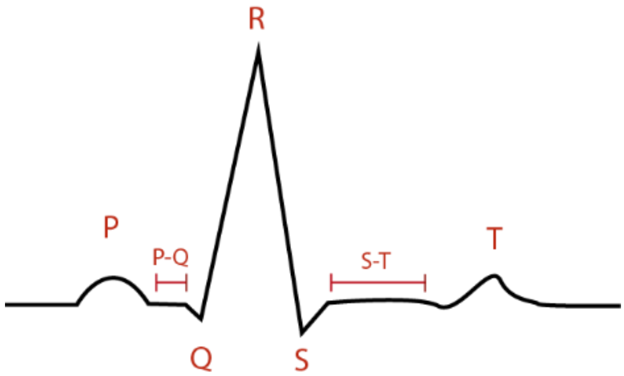



*Figure 1.1.* The major components of the cardiac cycle. When blood fills the atria, the sinoatrial (SA) node evokes electrical potentials, which causes the atria to depolarise. The P-wave represents the depolarisation of the atria. The electrical signals travel from the SA node to the atrioventricular node, which is represented as the P-Q segment. Then depolarisation of the ventricles occurs: the Q wave corresponds to depolarisation of the interventricular septum; the R-wave is produced by depolarisation of the main mass of the ventricles; and the S-wave represents the last phase of ventricular depolarisation at the base of the heart. The R-wave is the most dominant wave of the ECG. The S-T segment reflects the plateau in the myocardial action potential, and represents the ventricles contracting and pumping blood. Ventricular repolarisation occurs, represented as the T-wave.

Empirical research has indicated that heart rate increases with task difficulty (e.g. Sosnowski et al., 2004; Boutcher & Boutcher, 2006), unexpected situations (e.g. Schmitz et al., 2012), and with both cognitive and psychological stress (e.g. Jayasinghe et al., 2014; Wetherell & Carter, 2014). Heart rate has also been used to distinguish workload conditions during naturalistic sustained attention tasks. For example, Mehler et al. (2012) found that heart rate increased with cognitive workload during a working memory task while driving, although performance measures of steering wheel reversal rate and acceleration were unable to differentiate between different working memory demands. These results demonstrate the potential sensitivity of heart rate as a measure of attentional load.

HRV has increasingly been studied in relation to behaviour and cognitive processing as the variability in RR intervals allows for adaption to changing environment and task demands. Theoretical models suggest higher vagal tone represents greater adaptivity and is therefore often associated with better cognitive performance (e.g. Lehrer, 2013; Thayer et al., 2009). Hughes et al.'s (2019) meta-analysis found that moderate-to-strong changes in heart rate and RMSSD were sensitive to workload manipulations. In addition, a robust time-on-task effect of HRV has been demonstrated by numerous studies. Fairclough & Houston (2004) found that low-frequency HRV was sensitive to time-on-task and increased over time. Luque-Casado et al. (2013) found that time-domain measures of HRV (RMSSD, SDSD) decreased over time. The authors found that the largest reduction was found during a psychomotor vigilance task, compared to a cognitive control and perceptual task. Luque-Casado et al. (2016) also demonstrated a decrease in HRV parameters as a function of time-on-task during a vigilance task. These results point to the sensitivity of HRV to the demands of sustained attention, more than other task-related cognitive components such as cognitive control.

### 1.2.2 Electrodermal activity (EDA)

Another commonly employed measure of the ANS is electrodermal activity (EDA). Small amounts of sweat pass to the upper layer of the skin, the stratum corneum, via eccrine sweat glands depending on activity of the SNS. As sweat fills the duct, there is an increase in conduction through the resistant corneum: the more sweat, the lower the resistance of the ducts (Dawson et al., 2007). The changes in the value of the ducts result in EDA changes. EDA can be measured by placing a small current through a pair of bipolar electrodes placed on the skin, and so EDA measures changes in the skin's ability to conduct electricity depending on the activity of the sweat glands. As the eccrine sweat glands are solely under control of the SNS and has no relationship with the PNS, EDA is a simple and non-invasive peripheral index of the SNS (Boucsein et al., 2012).

The signal of EDA is typically decomposed into two time-domain measures: skin conductance level (SCL) and skin conductance response (SCR; Boucsein et al., 2012; see Figure 1.2). A large body of the literature focuses on the frequency and amplitude of SCRs, characterising rapid phasic activity. The SCR is often described as “peaks” of activity following stimuli but can also occur arbitrarily. Non-specific SCRs are responses not directly related to the onset of a stimulus. The latency of the SCR, the SCR half-recovery time (time between the SCR peak and 50% recovery of SCR amplitude), and SCR rise time (time between the onset and peak of SCR), can also be quantified (Andreassi, 2006). For research investigating states of arousal, SCL amplitude is commonly calculated, which demonstrates characteristics such as the slow climbing and falling of EDA over time. The SCL generates a constantly moving baseline, indicating that the SCL regularly changes within an individual, and so it is important to understand whether the differences in SCL are related to the experimental condition or related to within-subject variance. As SCL values are contaminated by SCRs, the minimum signal SCL value can be extracted, as the minimum value will always lie outside of an SCR. This method facilitates a truer representation of tonic changes and has been reported in previous studies (e.g. Braithwaite et al., 2014). Moreover, it is possible to analyse spectral characteristics, exploring low- and high-frequency components, albeit this is less common (Posada-Quintero et al., 2016).

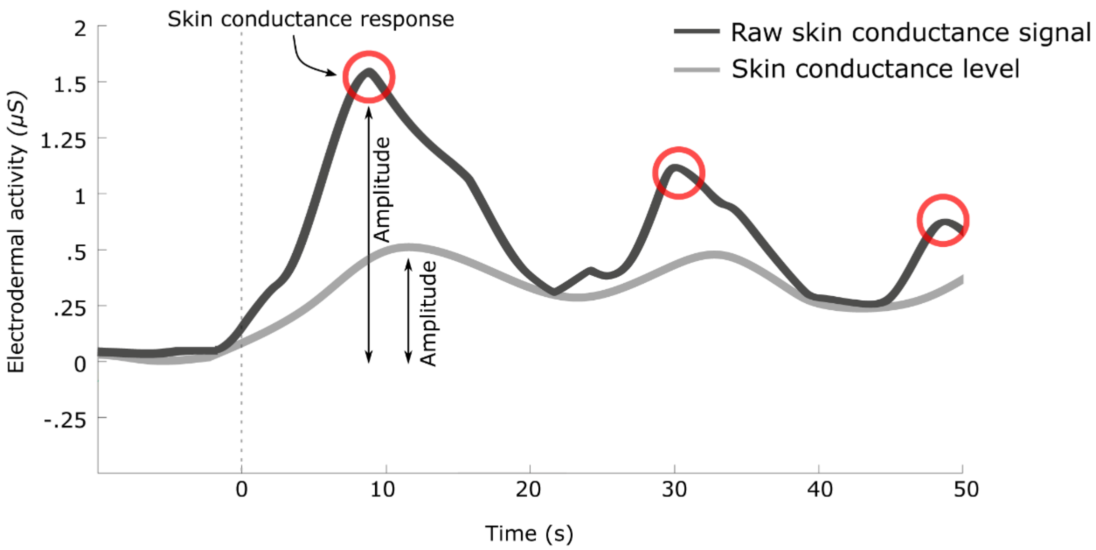



*Figure 1.2.* Skin conductance level (SCL) and skin conductance response (SCR). The dashed line at time point 0 s represents stimulus onset. Skin conductance responses appear as peaks in the raw data following stimulus onset, indicated by the red circles. Amplitude and frequency of the responses are often collected. Amplitude of the overall skin conductance level can also be computed, and the minimum skin conductance level signal can be extracted to ensure it is not contaminated by skin conductance responses.

Increases in sympathetic arousal resulting in increased SCL amplitude, SCR amplitude, and SCR frequency, are elicited during sustained tasks of continual performance when compared to resting levels (e.g. Andreassi, 2006; Dawson et al., 2007). Increased EDA is also associated with affective stimuli (e.g. Bradley & Lang, 2002), stressful situations (e.g. Wulvik et al., 2020), and higher mental demands (e.g. Visnovcova et al., 2016). Other studies have confirmed these findings during semi-naturalistic sustained attention tasks: EDA activity manifests as increased SCL and SCR amplitude during difficult driving (e.g. Schneegass et al., 2013) and dual-task driving (e.g. Ruscio et al., 2017). While some argue that increased EDA may represent allocation of attentional resources, others suggest that the changes are associated with affect and social interactions, as laboratory tasks are challenging stressors with social consequences (i.e. social desirability) resulting in increased sympathetic arousal (Dawson et al., 2007). Nonetheless, tonic and phasic changes in EDA are useful measures for investigating general states of arousal and alertness, and the response to novel stimuli, respectively.

### 1.2.3 Electromyography (EMG)

Electromyography (EMG) reflects the collective electrical signal from muscles during contraction and relaxation (Reaz et al., 2006). Surface EMG records muscle activity from multiple motor units from above the muscle on the skin. The composition of the motor unit, the number of muscle fibres per motor unit, the metabolic type of muscle fibres, and many other factors affect the shape of the motor unit action potentials in the EMG. Two parameters are commonly calculated: the root-mean-square (RMS) value and the average rectified value. Both provide useful measurements of the signal amplitude representing motor unit action potentials during contraction.

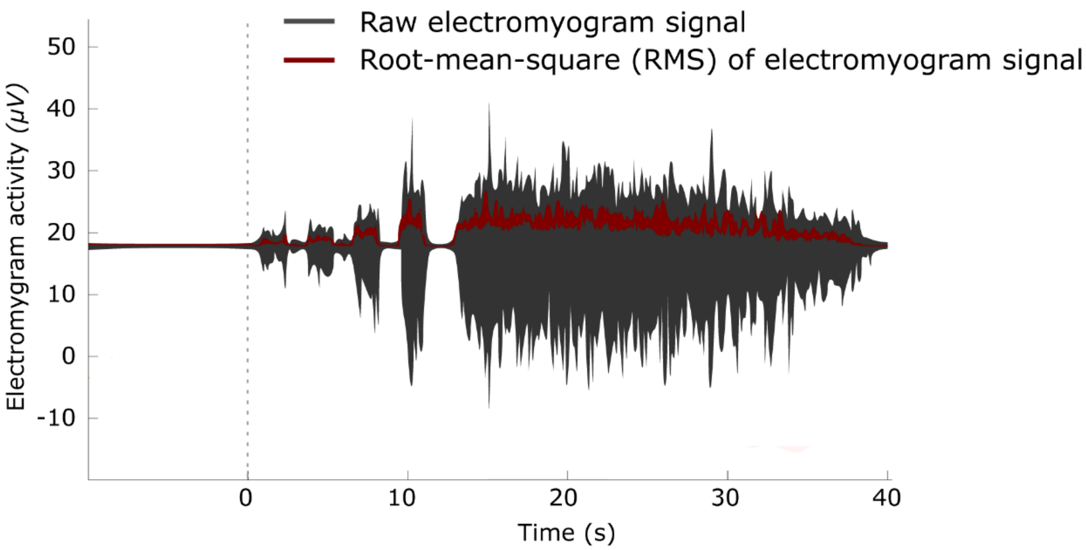



*Figure 1.3.* Electromyogram (EMG) signal. An example graph depicting the root-mean-square (RMS) envelope (red) overlaying the raw EMG signal (black). Dashed line represents stimulus onset.

Although commonly the amplitude of the EMG signal is used to differentiate between physical load, it can also be used to differentiate cognitive load. Facial EMG recordings can be obtained by using small surface electrodes over muscles that are associated with facial expressions, such as the corrugator supercilia muscle and frontalis muscle (e.g. forehead wrinkles). Sustained attention tasks, both externally paced and self-paced, have been associated with increases in facial EMG (e.g. Waterink & van Boxtel, 1994). However, electrodes on the face increase awareness of facial expressiveness, potentially leading to exaggerated or suppressed facial expressions. Less intrusively, muscle activity has been measured in the shoulder area and has been associated with cognitive and affective state.

Increases in muscle activity of the deltoideus and trapezius muscle have been linked to greater mental effort (e.g. Wijsman et al., 2013), stressful situations (e.g. Wijsman et al., 2010), and increased attentional demands (e.g. Waersted & Westgaard, 1996). Roman-Liu et al., (2013) found that trapezius and deltoideus activity increased similarly during a sustained attention task and a vigilance task, when compared to a control task. The demands of both tasks on muscle activity were similar. The authors argue that this reflects increased mental effort as body posture was the same during all tasks and the upper limbs were not impacted by button presses. Attempting to uncover the mechanisms, Wixted & O' Sullivan (2018) revealed a significant moderated (end-tidal $CO_2$) mediation (parasympathetic high-frequency HRV) model to explain the relationship between upper trapezius muscle activity and attention demands. They argue that sustained attention acts as a psychosocial stressor which increases upper trapezius muscle activity by inhibiting PNS activity and reducing end-tidal $CO_2$ increasing hyperventilation.

### 1.2.4 Electroencephalography (EEG)

As numerous functionally separable brain networks are involved in sustained attention (see Chapter 1.1.3), electroencephalography (EEG) has been utilised to uncover the mechanisms associated with attentional lapses. EEG measures the synchronised activity of post-synaptic electrical potentials at millisecond resolution and is an important method for studying the transient dynamics of the brain's large-scale neuronal circuits (Niedermeyer & Lopes da Silva, 2005).

A continuing trend in neuroimaging is the estimation of the mean level of spatial and temporal activation that is associated with experimental conditions, and as a result, EEG is often analysed by traditional univariate approaches. These univariate analyses can analyse EEG data in the time- or frequency-domain. In the time-domain, event-related potentials (ERPs) represent voltage fluctuations that are time-locked to an event, such as the onset of a visual, auditory, or somatosensory stimulus (Kappenman & Luck, 2012). ERP waveforms reflect the summation of thousands of post-synaptic potentials and can vary in amplitude, latency, duration, and topography (see Figure 1.4). As the signal-to-noise ratio and the waveform is so small, most ERPs only become visible when multiple EEG epochs are averaged together. Distinct ERP components that reflect specific neurocognitive processes have been associated with different waveforms. ERPs are often described either as early components, also termed sensory or exogenous, as they appear roughly within the first 100 ms or so post-stimulus; or late components, also termed cognitive or endogenous, thought to reflect evaluation processes (Sur & Sinha, 2009). The P100, an early sensory component, presents as a positive deflection at approximately 100 ms post-stimulus. For visual stimuli, it is detected over lateral posterior and occipital electrodes and is sensitive to the physical properties of the stimulus. However, research has also demonstrated that the P100 might be modulated by higher-level processing such as attention (Yamada et al., 2015). Another prominent ERP component is a large positive component that peaks at approximately 300 ms (though has been characterised as up to 900 ms) post-stimulus, and is termed the P300 (Linden, 2005). The P300 consists of two components. A fronto-central component termed the P3a, elicited by deviant stimuli; and a posterior-midline component termed the P3b, thought to reflect information processing including attention and working memory mechanisms. However, the underlying functional significance is still highly debated, and a recent review suggests that the P3b most likely characterises stimulus-response link reactivation (Verleger, 2020).

While ERPs can provide useful insights into the timing of neuronal events that subserve sensory, perceptual, and cognitive processes, each frequency present in the data can be decomposed to provide further insight into the parallel processes involved (Roach & Mathalon, 2008). Time-frequency analysis involves decomposing the EEG signal into magnitude and phase information for each frequency over time relevant to a stimulus of interest. Spectral composition of the data results in frequencies often characterised as: delta (1–4 Hz), theta (4–7 Hz), alpha (8–12 Hz), beta (13–30 Hz), and gamma (30+ Hz) bands. Analyses of the frequencies can tell us which have the most power, and how their phase angles synchronise in time and space. Changes in EEG power are thought to reflect underlying changes in neuronal synchrony: the greater the power, the more neurons firing together. In many cases, including during sustained attention tasks, EEG activity cannot be assumed to be stable over time. Therefore, the event-related spectral perturbation (ERSP) technique can be used to characterise event-related changes in power with methods such as the Wavelet transform or the Hilbert transform, relative to a pre-event baseline period, in epochs time-locked to a stimulus. Baseline normalisation procedures, such as decibel (dB) conversion, are also advantageous as background and unrelated task activity are removed (Cohen, 2014). Squaring the magnitude values and then averaging over trials results in a 2-dimensional matrix, a spectrogram, containing total power of the EEG at each frequency and time point (see Figure 1.4). The total power comprises both evoked and induced neural power. Evoked power reflects EEG data that has phase-locked to the event across trials, whereas induced power refers to power that is not phase-locked. Mean power is calculated and is reflected as a positive or negative number relative to the baseline period. Event-related desynchronisation (ERD) refers to localised amplitude attenuation of rhythms; event-related synchronisation (ERS) refers to amplitude enhancement. The ERSP approach has been employed in a variety of studies covering a broad range of different cognitive task demands, including sustained attention and vigilance tasks.

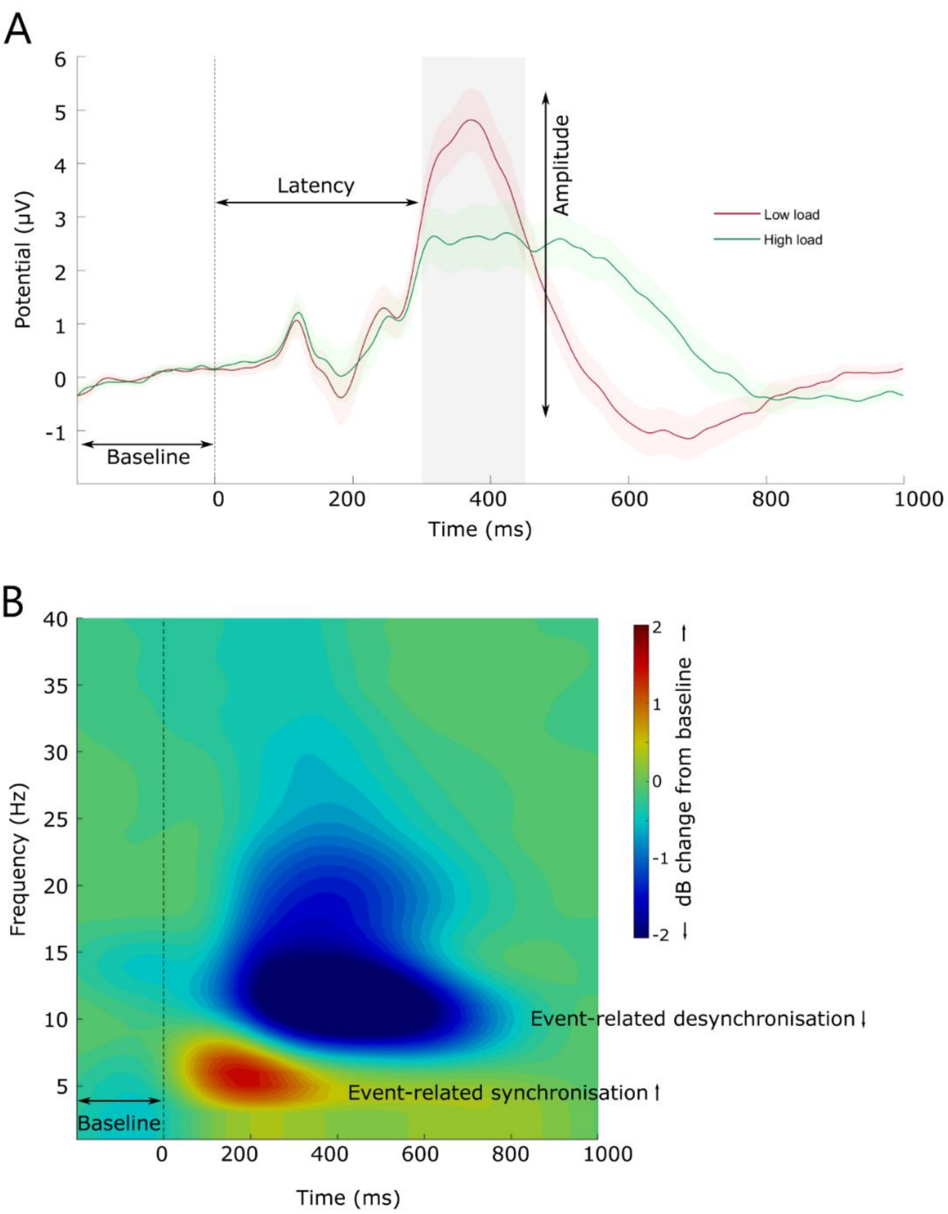



*Figure 1.4.* Univariate analyses of EEG data. **(A)** Event-related potential (ERP) waveform. ERP waveforms reflect the summation of thousands of post-synaptic potentials and can vary in amplitude, latency, duration, and topography. Visualisation of ERP waveforms often represents the differences by superimposing the conditions of interest. In this example, a low load condition overlaps a high load condition. Importantly, no difference in amplitude is apparent during the baseline period, validating the post-stimulus differences. Shaded areas represent ± standard error of the mean difference. Dashed line represents onset of the stimulus. **(B)** Time-frequency spectrogram. The event-related spectral perturbation (ERSP) technique is used to characterise event-related changes in EEG power. The ERSP approach results in a 2-dimensional matrix containing mean power at each frequency and time point. Mean power is reflected as a positive or negative number relative to the baseline period. Event-related desynchronisation (ERD) refers to localised amplitude attenuation of rhythms; event-related synchronisation (ERS) refers to amplitude enhancement. Dashed line represents onset of the stimulus.

Neural oscillations have been implicated in three keys aspects of sustained attention: monitoring ongoing processes, activation of task-relevant processes, and inhibition of task-irrelevant processes (Clayton et al., 2015; Stuss et al., 1995). Fronto-medial theta activity has been implicated in the cognitive control and monitoring of ongoing processes (for reviews, see Cavanagh & Frank, 2014; Sauseng et al., 2010). In general, fronto-medial theta power increases in combination with deteriorating performance during sustained attention tasks (e.g. Wascher et al., 2014). In particular, theta power demonstrates a continuous increase with time-on-task, thought to reflect increased effort (Wascher et al., 2014). Contradictory research, however, suggests that fronto-medial theta power can be positively associated with effective attentional control. For example, Cooper et al. (2017) found increased fronto-medial theta power with reduced behavioural variability (i.e., increased efficiency of cognitive control) during task switching. Clayton et al. (2015) suggest that this discrepancy could be due to the timescale of the tasks. Theta power reflects the engagement of cognitive control processes, which during short tasks, improves performance. During sustained tasks when cognitive resources are exhausted, performance reduces. Therefore, over time, an increase in theta power is associated with a decline in performance. The second crucial functional role of cortical oscillations concerns the activation of task-relevant processes. High-frequency gamma oscillations have been implicated in enhancing attention to sensory inputs, and reliably appear in task-relevant areas (for a review, see Clayton et al., 2015). For example, Friese et al. (2016) found increased gamma activity in auditory and visual cortices during a bimodal attention task; and Proskovec et al. (2019) found a positive correlation between gamma oscillations within the anterior cingulate and task-switching performance – the anterior cingulate thought to be part of a critical network involved in selecting relevant external stimuli and internal goals (Menon & Uddin, 2010). In addition, alpha ERD (i.e. reduced alpha power compared to baseline) has been associated with cortical activation (Pfurtscheller, 2006). Alpha ERD is often found in parietal-occipital areas, thought to be associated with the executive network (Petersen & Posner, 2012; Posner & Petersen, 1990; see Chapter 1.1.3). Reduced alpha ERD (i.e. higher alpha power in the condition of interest but still suppressed compared to baseline) has been related to poorer performance (e.g. Dimitrijevic et al., 2017; Proskovec et al., 2019). Finally, a robust correlate of inhibition of task-irrelevant areas is alpha ERS (Klimesch, 2012). Where alpha ERD is thought to represent active cortical processing, alpha ERS is thought to reflect inhibition (Klimesch, 2012). Foxe et al. (1998) found increased alpha power over visual cortices when attention was directed towards the auditory element of a bimodal sensory stimulus. Therefore, increases in alpha power reflect attenuation of irrelevant information and distractors, supporting attentional processes. In summary, fronto-medial theta ERS is thought to represent cognitive control processes, gamma ERS and alpha ERD indicates cortical processing and task engagement, and alpha ERS is associated with the inhibition of task-irrelevant areas.

### 1.2.5 Eye tracking

Many of the pioneering studies in attention research have investigated visual attention centred on ocular behaviours. Eye blinks, fixations, saccades, and pupil diameter have all been used to measure attention control and alterations in arousal state. The fixation point can be considered the site of overt attention. A fixation is composed of slow and minute movements (microsaccades, tremor and drift) that help the eye align with the target and avoid perceptual fading (Martinez-Conde et al., 2004). Average fixation durations typically vary between 150 – 300 ms (Tullis & Albert, 2013), however longer fixations have been reported (e.g. Otero-Millan et al., 2008). In basic visual processing research, fixation duration increases with visual scene complexity (e.g. Pomplun et al., 2013), increases with cognitive load (e.g. Rayner, 1998), decreases with positive affect (e.g. Tichon et al., 2014), and increases under uncertain situations (Brunyé & Gardony, 2017).

Pupillometry is the measurement of pupil size and reactivity. Visual attention allocation and mind wandering can influence the size of the pupil (Pelagatti et al., 2018). The mechanisms behind this association reflects fluctuations in neuronal firing and the release of noradrenaline in the locus coeruleus, which has been associated with subtypes of attention and arousal (see Chapter 1.1.2 and 1.1.3; Aston-Jones & Cohen, 2005; Joshi et al., 2016). Mind wandering states have been associated with smaller diameters (Unsworth & Robison, 2018), distracted states have been associated with larger pupil diameters (Unsworth & Robison, 2016), and both states have been associated with reduced performance (Unsworth & Robison, 2016, 2018). Therefore, measures of fixation and pupil diameter can index attentional shifts and demands, as well as general arousal state.

## 1.3 Applications for autonomic and neurophysiological monitoring

Integrating models to better understand performance decrements in sustained attention have implications for understanding attentional and cognitive dysfunctions across a wide range of applied contexts. Decrements can be disastrous during human-machine interactions in everyday and occupational settings that require sustained attention. Studies have extensively demonstrated that impairments in sustained attention are associated with driving errors (e.g. Schmidt et al., 2009), medical errors (e.g. Taylor-Phillips et al., 2014), and pilot errors (e.g. Endsley & Garland, 2000). Poor sustained attention is also associated with many clinical disorders including attention deficit hyperactivity disorder (e.g. Gmehlin et al., 2016), autism (e.g. Chien et al., 2014), and schizophrenia (e.g. Giakoumaki et al., 2011). Attentional deficits during human-machine interactions and clinical populations, and the potential for autonomic and neurophysiological monitoring, are discussed in more detail below.

### 1.3.1 Human-machine interaction

Human-machine interaction refers to the communication and interaction between a human and a machine normally via a form of interface. Occupational settings such as manufacturing (e.g. Cascio & Montealegre, 2016), medical (e.g. Bologva et al., 2016), and aviation (e.g. Canino-Rodríguez et al., 2015) are commonly cited as workplaces that rely on successful human-machine interaction. One of the main aspects of interaction concerns working with automation safely and effectively. Automated systems and digital information are commonly found in industry, particularly in the transport, manufacturing, and medical industries. Rather than the human manually carrying out laborious tasks, the role of the human changes, becoming responsible for monitoring the automation and working processes, and intervening when automation fails (Kassner et al., 2017).

Although automation is intended to reduce human errors, research has indicated that monitoring automation can result in poor performance during periods of low situation awareness due to high system complexity and lack of expertise (Körner et al., 2019). Therefore, attentional lapses are common and come with severe consequences. Insufficient monitoring of an automated system and the environment is a contributing factor in major transportation incidents. A widely cited example is the cruise ship Royal Majesty (e.g. Parasuraman & Manzey, 2010; Parasuraman & Riley, 1997). The navigation system lost the GPS signal due to physical damage to the automatic radar plotting aid. The ship steered towards shallow waters and became grounded. This is an example of an incident due to a failure of sustained monitoring of an automated system. Scenarios such as these have found that incidents occur due to an insufficient response by the operator due to complacency and overreliance of the system, resulting in delayed responses or a failure to act altogether (Parasuraman & Manzey, 2010). Deficits in attentional allocation play a critical role in the misuse of automated systems as attention allocation favours other tasks over automation monitoring, particularly when users over-trust the system (Parasuraman & Manzey, 2010) or lack understanding of the system (Choi & Ji, 2015).

Everyday sustained attention tasks are also under continual threat of lapses in attention. Driving is a common task where momentary lapses can have life-threatening consequences. The driver integrates multiple streams of information from the driving environment, engaging a range of perceptual, cognitive, and motor skills such as spatial awareness, motor function, and navigation. Humans shift their attention between the multiple requirements while ignoring irrelevant sensory distractors. Driving research has indicated that lapses in attention associated with traffic accidents have been related to a range of distractions including mobile phones (e.g. Strayer & Drews, 2007), in-vehicle systems (e.g. Arexis et al., 2017), and eating (e.g. Tay & Knowles, 2004). Predictions that autonomous vehicles may be able to prevent accidents have led to scientific advancements in engineering in both industry and academia. The Society of Automotive Engineering illustrated six levels of automation, ranging from 0 "No automation", to 5 "Full automation" (SAE, 2018). While Levels 2 and 3 require a driver to monitor the environment and take back control of the vehicle when requested; Levels 4 and 5 require no input from the driver. As human errors and judgment are likely to be eliminated from autonomous driving situations, it is argued that autonomous vehicles could reduce the accident rate and fatalities. While vehicle automation holds the promise to improve traffic safety, distinct and novel opportunities for human error are introduced. For example, autonomous vehicle technology must be programmed by humans for specific rule-following behaviour, indicating that vehicle automation is still initially reliant on human judgement and decision-making. In addition, transportation research as discussed above has demonstrated that rather than the human being eliminated, the role of the human changes: humans must adapt to the role of monitoring automation. Therefore, monitoring vehicle automation during Level 2 and 3 driving may lead to negative consequences associated with the misuse and disuse of automated systems (Parasuraman & Riley, 1997). In addition, a remote operator will need to monitor automation for Level 4 and 5 vehicles. Because of this, it has been argued that assessing whether a driver's functional state is suitable during semi-autonomous driving is crucial (Collet & Musicant, 2019), as well as the remote operator during fully autonomous driving. A driver state monitoring (DSM) system, including cognitive and affective indices, may improve safety during autonomous driving. A DSM system continuously monitors a user using hybrid measures including biological (e.g. muscle activity) and physical measures (e.g. blink frequency). By synthesising and classifying functional state, the system could provide feedback to the passenger or adapt vehicle behaviour (see Figure 1.5). DSM systems are discussed further in Chapter 6.2.

### 1.3.2 Clinical populations

Understanding performance decrements related to attentional and cognitive dysfunction will provide a better understanding of clinical disorders and improvements in interventions. Lasting sustained attention deficits are pervasive in a broad range of clinical populations including developmental (e.g. attention deficit hyperactivity disorder, Gmehlin et al., 2016), psychiatric (e.g. schizophrenia, Giakoumaki et al., 2011), and neurodegenerative disorders (e.g. Parkinson's disease, Zgaljardic et al., 2003). Currently, there is no universal test for assessing sustained attentional deficits in clinical populations (Fortenbaugh et al., 2017). In addition, enhancing sustained attention has important implications for improving clinical outcomes. Localising the brain mechanisms responsible for sustained attention can improve the specificity of pharmacological interventions (Davidson, 2008) and brain stimulation techniques (Nelson et al., 2014). Moreover, novel training methods can be used to influence underlying neural networks. As sustained attention shares a significant relationship with other cognitive functions, cognitive training targeting sustained attention should improve several higher cognitive functions.

Two forms of training are most widely used in the literature: meditation and computer-based cognitive training. Forms of meditation have been shown to improve attention and reduce negative psychological experiences (Tang et al., 2007, 2012, 2015). In particular, focused attention meditation has shown to reduce activation of the default mode network (Brefczynski-Lewis et al., 2007), as well as improve performance on a sustained attention task (MacLean et al., 2010). Moreover, training programmes on sustained attention tasks have demonstrated improved executive attention function (DeGutis et al., 2017; Van Vleet et al., 2016). DeGutis et al. (2017) found that patients with Parkinson's disease improved in a visual search task after one-month (7.3 hours) of tonic and phasic attention training (a continuous performance task). These results provide evidence that sustained attention training can reduce spatial biases. Age-related cognitive decline of sustained attention has also been targeted. Van Vleet et al. (2016) found that two weeks of tonic and phasic attention training on older adults improved working memory and verbal fluency. This also supports the view that cognitive training can improve a range of executive functions.

Brain-computer interface (BCI) technology enables a direct communication pathway between the human brain to an external device (see Figure 1.5). Using sensors and algorithms, the BCI analyses neural signals, extracting relevant patterns and adapting them to control the device. The development of BCIs were initially focused on assisting rehabilitation of paralysed individuals by replacing defective sensory input or providing a way of interacting. Recently, several new applications for BCIs have been proposed, including to improve cognitive functioning. A cognitive BCI can monitor fluctuations in sustained attention in real-time. Gaume et al. (2019) recently developed a cognitive BCI and found that prefrontal theta, gamma, fronto-central beta and alpha were able to distinguish between task difficulty (easy versus medium/hard) of a continuous performance task and provided a classification accuracy of 85% for 30 s epochs. The BCI could then warn the user immediately if their attention levels drop or provide continuous feedback to train sustained attention abilities. This is also the primary principle of neurofeedback, which relies on the rapid decoding of brain state from real-time neural oscillatory data to provide participants with feedback on a moment-to-moment basis about activity within a predetermined brain region (e.g. prefrontal cortex). Using mental imagery, participants upregulate or downregulate brain activity. Despite that BCIs rely solely on decoding neural activity, considering the relationship between sustained attention and arousal, decoding signals of the autonomic nervous system, biofeedback, may also be a suitable method for improving deficits in sustained attention. In summary, uncovering the mechanisms relatable to sustained attention could enable successful neuro- and biofeedback for healthy and clinical populations and development towards advanced BCI and biofeedback systems.

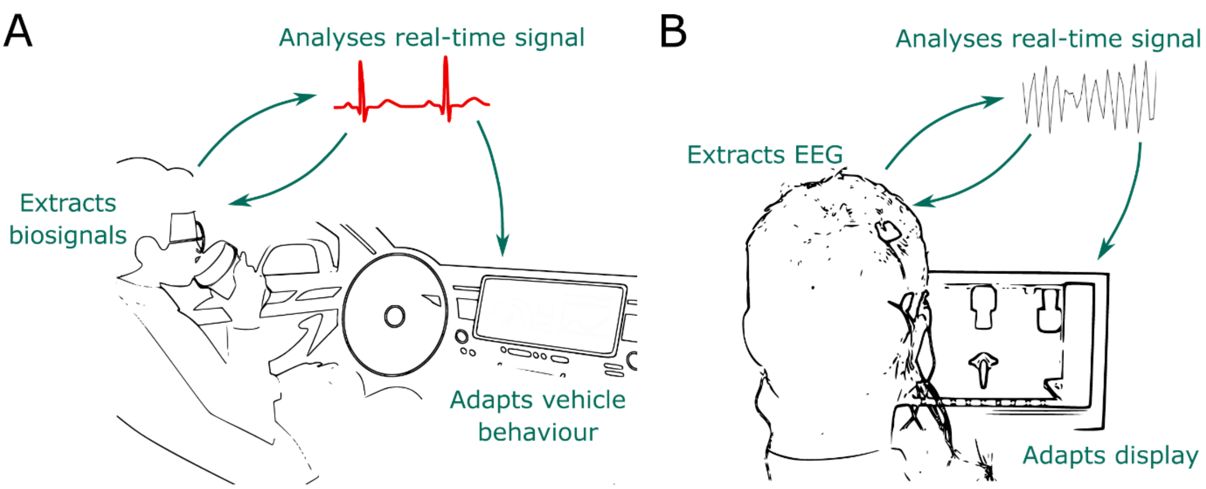



*Figure 1.5.* Applications for autonomic and neurophysiological monitoring. **(A)** Driver state monitoring during autonomous driving. **(B)** Neurofeedback for cognitive training.

### 1.3.3 Applied and laboratory-based research

Despite the relevance of attentional lapses during everyday sustained tasks, most of the literature is based upon simple unnaturalistic tasks in strictly controlled laboratory-based settings. While basic research is important for understanding the psychophysiological and neural mechanisms underlying attention, it is difficult to provide practical guidelines for measuring attention in naturalistic settings, such as occupational or driving contexts. Measuring sustained attention separately from other cognitive behaviours and processes can make it difficult to transfer the results to complex naturalistic settings where cognitive behaviours and processes overlap, and there are several methodological limitations concerning laboratory-based attention tasks. Moreover, continual monitoring of human state must be able to capture accurate and reliable physiological signals in naturalistic settings where external artefacts will likely impact the measurement. These concerns are further discussed below.

The models of sustained attention discussed in Chapter 1.1.2 demonstrate that sustained attention relies on a large-scale network of brain areas that overlap with a variety of higher cognitive functions including other aspects of attention and executive control. It is acknowledged that rather than being fully independent from one another, often these systems are activated together in everyday situations. For example, the alerting and orienting system (Petersen & Posner, 2012; Posner & Petersen, 1990) are both activated when an event informs where and when a target occurs (Fan et al., 2009). A further element of this considers overt and covert shifts of attention. Over the last 50 years, cognitive neuroscience research has typically focused on covert orienting of attention for visual scene processing: participants attend to an extra-foveal region of space without a change of fixation. However, in everyday contexts, attention is often coupled with eye movements. For example, ocular behaviour and motor execution are intrinsically linked both spatially and temporally, and active drivers successfully fixate directly at the objects being interacted with or ones that precede the action. In addition, covert attention requires processes involved in the inhibition of the saccadic response, therefore activates processes that are likely unemployed in everyday settings. The frontal eye fields play a critical role in the allocation of spatial attention and have been shown to interact with attention-related parietal alpha (Capotosto et al., 2009). Consequently, it is important to adopt overt shifts of attention to improve the transferability of the results to applied contexts, which include interactions between ocular and cognitive behaviours under naturalistic conditions.

Several limitations with laboratory-based research include the attentional task itself. For example, Ho et al. (2014) argue that laboratory-based paradigms are too short in duration, typically less than an hour, and therefore it is difficult to infer from the findings about sustained attention in naturalistic settings. These scenarios are unable to model the physical and mental state of the user in real scenarios. Conditions are often low arousal and low stress, and warning signals or experimental instructions may be semantically meaningless. In addition, participants often partake in a quiet and isolated room, where varied distractions are limited. This differs from naturalistic settings, where multimodal distractions dominate, lapses in attention may result in significant, potentially life-threatening consequences, and these scenarios are often associated with high arousal states. Therefore, applied research is needed to understand the impact of such overlapping variables on sustained attention.

A final concern regards the practical and logistical limitations of applied monitoring of autonomic and neurophysiological indices verified in the laboratory. Traditional methods for ECG, EDA, EMG, and EEG acquisition typically rely on wet silver/silver chloride (Ag/AgCl) transducing electrodes. These electrodes convert ionic current to electronic current for amplification and subsequent signal processing. Abrasion of the skin and a conducting gel between the electrode and the skin is required to improve electrical impedance. However, wired systems using wet electrodes can be intrusive and disorderly. The awareness of being measured can have a negative effect on the participant, subsequently confounding physiological measurements and reducing the ecological validity of the findings (Oken et al., 2015). Crucially, wired systems are restricted to a laboratory setting and are impractical for applied studies. For autonomic and neurophysiological monitoring to be successful, unobtrusive continuous physiological monitoring devices are needed (Alberdi et al., 2016). However, artefacts arising from muscle movement, eye movement, environmental noise and other complications are amplified in applied environments. For example, regarding pupillometry, it is difficult to collect reliable pupil diameter estimates in naturalistic settings. Large variations in luminance found in naturalistic studies will create noise in the pupillary signal (Lohani et al., 2019). In addition, movement in eye position will shift the pupil from a circular to elliptical shape, impacting pupil size (Gagl et al., 2011). Therefore, targets of interest may record a change in pupil size irrespective of actual changes in dilation. In summary, these limitations must be considered before utilising such measures in applied research.

Basic research has the potential to guide the development of detecting attention lapses in the real world. However, in naturalistic environments, the distinct autonomic and central cortical mechanisms related to different aspects of attention may be considered less important (Sarter et al., 2001). Tasks are rarely performed without impacting other aspects of executive functions and cognitive processes, and attentional networks are to some extent intertwined and often overlap in activation (Sarter et al., 2001). Similarly, in dynamic scenarios, aspects of attention such as perceptual load or cognitive load will not be the only contributing factor to attentional deficits (Murphy et al., 2016). Applied research is therefore a promising means for investigating attention, yet limitations in physiological measures in naturalistic environments must be considered before undertaking applied studies.

## 1.4 Thesis overview

In this introductory Chapter, I have argued that sustained attention is a multifaceted concept that operates on a series of timescales, impacted by global mechanisms of both vigilance and arousal, as well as being a key element of many higher-order cognitive functions. Considering that lapses in attention fluctuate over time, I argued that continual monitoring of autonomic and neurophysiological indices with good temporal resolution may reveal the mechanisms that underlie attentional lapses, as well as provide preventative measures by assessing when they are more likely to occur. I then introduced several physiological and neurophysiological measures that have been used in the literature to capture attention over time. The final part of the Chapter revealed the relevance of sustained attention to everyday tasks such as driving, occupational settings such as monitoring automation, and clinical populations such as Parkinson's disease. I also argued that measuring and uncovering the mechanisms of sustained attention during naturalistic studies are required, as everyday attention is impacted by dynamic and overlapping cognitive behaviours and processes.

The aims of the remainder of the thesis are threefold. Firstly, I aim to understand the lasting effects of a previous task on autonomic arousal and neural mechanisms related to sustained attention (**Chapters 2, 3, and 4**). This aims to provide novel research uncovering the neurobehavioural effects of task switching which is a common requirement in everyday activities. This thesis will focus on a semi-autonomous driving scenario when the passive driver is required to 'takeover' from the automated driving system. Does the attentional demand of an unrelated task have an impact on takeover performance and processes? Secondly, I aim to understand how attention fluctuations manifest during a naturalistic study of automated driving during an unexpected event, and whether it is possible to monitor autonomic mechanisms accompanied by a variety of real-world artefacts (**Chapter 5**). This is to understand how attention fluctuates during autonomous driving, and whether these fluctuations can be measured in an applied context. Finally, I attempt to unravel the acceptability of a biofeedback system during simulated automated driving, and whether functional state can be captured as changes in visual attention or autonomic arousal (**Chapter 6**). This is to understand whether autonomic monitoring during sustained attention negate the potential benefits resulting in visual distraction and overload.

# 2.0 Physiological and Neurophysiological Measures of Low and High Workload During Multitasking

## 2.1 Overview

The General Introduction Chapter suggested that continual monitoring of autonomic and neurophysiological indices with good temporal resolution may be able to capture attentional lapses associated with poor performance, as well as provide preventative measures by assessing when they are more likely to occur in applied settings. To understand which measures of autonomic arousal and neural mechanisms related to sustained attention are impacted by a task of differing attentional load, the current Chapter reports the first experiment exploring different electrophysiological methods and their relationship with workload. Participants experienced cognitively demanding multitasking at low and high workload intensities. Physiological (ECG, EDA, EMG), neurophysiological (EEG), subjective (verbal ratings of 'stress'), and behavioural (task performance) data were collected.

## 2.2 Introduction

Humans frequently engage with multiple tasks in parallel. The office worker negotiates with a client on the telephone while typing an email to another client. The parent holds their distressed baby while cooking dinner and conversing with their partner. This need to manage more than one task is ubiquitous across work and at home, and so many researchers in recent years have investigated multitasking effectiveness in daily life. For example, media-multitasking is a fruitful area of research focusing on performing several digital tasks at once (Carrier et al., 2015) and has a significant impact on cognitive functioning (for a mini meta-analysis, see Wiradhany & Koerts, 2019). In more laboratory-based multitasking paradigms, robust findings demonstrate a significant multitasking cost to human performance. Reaction time and accuracy measures deteriorate when tasks are carried out simultaneously compared to in isolation (e.g. Glass & Kang, 2019; Lin et al., 2016; Poljac et al., 2009). Performance decrements are also found when a distinct task follows a previous task (Kiesel et al., 2010; Rogers & Monsell, 1995).

It is thought that multitasking deficits are due to information processing bottlenecks, suggesting that specific cognitive processes cannot be effectively carried out concurrently (e.g. Marois & Ivanoff, 2005; Pashler, 1984). Some researchers have argued that response selection is serial, which assumes that although cognitive resources coexist with one another, these are sequential in nature (Pashler, 1994). Therefore, only one resource can be used by a single task. For example, if two tasks require auditory perceptual mechanisms, only one task will be able to successfully engage auditory perceptual resources at a time. However, other researchers have demonstrated that parallel processing is possible (Hommel et al., 1998). Parallel processing models suggest that serial processing is still the most efficient and effective strategy. According to this view, if two tasks require auditory perceptual mechanisms, auditory perceptual mechanisms successfully engage in one task, and then the other, to increase performance efficiency. As such, information processing can be flexible and respond to environmental demands. Koch et al.'s (2018) integrated review emphasises the importance of cognitive flexibility and plasticity during multitasking, in addition to structural capacity limitations. For example, preparation before a switch can improve performance (cognitive flexibility; Vandierendonck et al., 2010) and practice can reduce performance costs (cognitive plasticity; Strobach & Schubert, 2017). As such, the bottleneck is not structurally implemented and is modulated based on the flexibility and plasticity of the cognitive system.

### 2.2.1 Multitasking paradigms

Multitasking is most typically investigated utilising dual-tasking and task-switching designs. Dual-task paradigms usually require participants to undertake two tasks simultaneously and in isolation. Performance deficits are found during dual tasking in comparison with single tasking (e.g. Zhang et al., 2018). Another variant examines dual-task interference by the temporal overlap of both tasks (the Psychological Refractory Period paradigm; Welford, 1952). The shorter the stimulus onset asynchrony (i.e. the greater the temporal overlap of task processing), the greater the performance deficits in the following task. Task-switching paradigms have also been utilised, where participants respond to two different tasks in a repeating or alternating sequence. Performance deficits are commonly found in the alternating trials (Rogers & Monsell, 1995). Both these paradigms are utilised in multitasking research as cognitive processes relating to two tasks overlap in time (Koch et al., 2018).

Studies investigating the neural mechanisms underlying task interference have driven our understanding of multitasking costs. For example, error processing is a mechanism that may interfere with multitasking performance. Weißbecker-Klaus et al. (2017) utilised a dual-task paradigm and measured event-related potentials (ERPs) associated with error processing. During dual-task conditions of greater complexity, the error-related negativity (Ne) component and the error-related positivity (Pe) component were delayed and Pe amplitude reduced, representing decrements in error detection and errors awareness, respectively. The post-error slowing component was diminished, suggesting participants did not adapt after an error. This provides evidence that error processing mechanisms were affected during multitasking. Ne and Pe ERP components are measured maximally over frontal-central sites suggesting the involvement of frontal mechanisms during multitasking. Other neuroimaging studies with superior spatial resolution have supported this view, associating fronto-parietal regions with multitasking costs. Al-Hashimi et al.'s (2015) functional magnetic resonance imaging (fMRI) study required participants to undertake a tracking task and a forced-choice discrimination task simultaneously and in isolation. Visual, frontal, and parietal regions were associated with multitasking. However, only the superior parietal lobule was associated with performance; increased activation positively correlated with performance. These results suggest that during multitasking, the superior parietal lobule is important for visuospatial attention, and an increase in activity is associated with improved task performance.

Although task switching and dual tasking arguably measure the same concept, they are typically studied independently from one another (Koch et al., 2018). However, their paradigms are distinct. For example, the stimuli presented in task-switching paradigms are related to both tasks. Participants may be presented with coloured animals as stimuli. Participants alternate between judging the colour (Task A) and the animal (Task B). Yet, the stimuli presented in dual-tasking paradigms are unique to the task. For instance, participants alternate between a visuomotor tracking task, a forced-choice discrimination task, and both tasks simultaneously. Ultimately, task switching focuses on studying how attention is reoriented from one task to another, whereas dual-tasking examines the impact of dividing attention. Overlapping fronto-parietal regions have been reliably activated for both types of tasks, despite the different processes involved. Ward et al. (2019) aimed to disentangle these processes by administering the paradigms in isolation and in combination. Dual-task activation was found in right fronto-parietal regions, whereas task switching elicited left fronto-parietal regions. In combination, activation was found in the left and right parietal cortices. As the authors note, it is difficult to infer distinct cognitive processes involved in dual tasking and task switching, as the fronto-parietal network is recruited during many processes involved in multitasking such as error processing, selective attention, and working memory (Harding et al., 2015).

### 2.2.2 Everyday multitasking

Dual-tasking and task-switching laboratory paradigms demonstrate that multitasking results in a performance cost implicating a broad-range of fronto-parietal regions. These methods offer high controllability and standardisation, but they often employ abstract tasks and simple stimuli.

Multitasking in real life consists of numerous context-relevant tasks, complex multisensory environments, and, at times, instructive feedback. Therefore, it is difficult to apply these findings to real-world environments. More ecologically valid paradigms can provide insights into multitasking which laboratory-based experiments cannot (Li et al., 2005).

Due to the occupational relationship between multitasking and cognitive workload, some researchers have adapted tasks used for pilot selection. Puma et al. (2018) utilised the Priority Management Task: an airline recruitment task that increases from managing one to four tasks simultaneously. Tasks such as gauge monitoring and a joystick tracking task utilise perceptual and motor pathways and are more typical to real-world scenarios, particularly in occupational settings. Measuring EEG, they found frontal theta and parietal alpha oscillatory activity increased in line with the number of tasks, for medium and high performers. Low performers exhibited high theta and alpha power throughout, indicating a ceiling effect. Although theta increases were as expected (see Chapter 1.2.4), alpha activity demonstrated the opposite anticipated effect, as lower alpha power is normally associated with higher task demands. During multitasking, the increased alpha could explain an increase in inhibition of irrelevant information (Klimesch et al., 2007), rather than deficits in attentional mechanisms.

Another multitasking platform utilised in the literature is the Multitasking Framework which administers up to four distinct tasks under low, medium, and high workload intensities (Wetherell & Sidgreaves, 2005). Unlike the Priority Management task that focuses on a particular work environment, the Multitasking Framework (see Figure 2.1) involves generic tasks requiring various time-limited responses, similar to many working environments. Wetherell & Carter (2014) administered the framework at low, medium, and high workload intensities. Participants participated in mental arithmetic, auditory monitoring, visual monitoring, and a Stroop task for 15 minutes. Measures of autonomic arousal including peak heart rate and systolic and diastolic blood pressure displayed greater values during the task compared to baseline. Perceived workload increased as workload intensity increased; in particular, mental demand and effort dimensions increased with each intensity. Wetherell et al. (2017) also found similar results. These findings demonstrate that multitasking has an impact on the reactivity of autonomic arousal during a more ecologically valid laboratory-based task.

### 2.2.3 Experiment rationale

The study aimed to understand which physiological signals provide a sensitive measure of attentional load on multitasking. In this study, several physiological measures were explored reflecting autonomic arousal (ECG, EDA, EMG), central mechanisms (EEG), and subjective ratings. Attentional load was altered via workload intensity and induced with the Multitasking Framework (low workload versus high workload). Physiology was measured before, during, and after the task, to measure tonic and phasic changes affected by workload induction.

This study differs from previous studies in three key respects. Previous research on multitasking focuses on a limited number of measures (e.g. Al-Hashimi et al., 2015; Puma et al., 2018). Considering that methodological differences are high between studies, this restricts understanding of the effectiveness of different measures in the same context. In addition, it is not clear which measures are most sensitive to workload, here being used to index attentional demand. Due to the complex nature of multitasking, it is often difficult to interpret the nature of the cognitive processes involved. Therefore, utilising a number of psychophysiological measures will provide a comprehensive dataset focusing on different aspects of biological mechanisms, and will provide an opportunity for well-informed interpretation of the data. Using continuous measures of human electrophysiology may improve understanding of time-varying processes under complex evolving environments. Secondly, it is important that laboratory multitasking involves several distinct tasks that capture cognitive functions in various processing streams, akin to naturalistic activities such as driving. The Multitasking Framework allows the inclusion of different tasks such as visual attention, psychomotor, and auditory processing. The tasks utilised, such as visual monitoring and data entry, are also comparable to tasks undertaken in some workplace environments. In addition, repeated multitasking does not result in habituation of performance (Wetherell & Sidgreaves, 2005). To the best of my knowledge, this is the first study that utilises the Multitasking Framework with a wide range of physiological measures to inform the interpretation of the data. Finally, numerous studies measure tonic changes in physiology during a baseline period and the task, however some studies (e.g. Marker et al., 2017; Salah et al., 2018) do not consider the physiological recovery period, which allows for investigation of changes in physiology from the task back to baseline (Laborde et al., 2017). In studies of static spectral analyses of neural data, continuous multitasking research typically does not statistically compare oscillatory activity from an event to a baseline and/or recovery period, or normalise the EEG data of interest to a pre-stimulus baseline (e.g. Puma et al., 2018). A recovery period enables the evaluation of trends across time; if the variable of interest increases activity over the task and the recovery period, then this could simply be a time effect rather than a true effect of the variable. As such, it is important to separate task-related activity from ongoing background activity, therefore including a baseline and a recovery task phase will enable this.

In accordance with past literature, autonomic arousal should increase during the task when compared to baseline and recovery, represented as increased heart rate, reduced heart rate variability time-domain measures, increased skin conductance level amplitude, increased frequency and amplitude of skin conductance responses, increased root-mean-square of muscle activity, and increased self-report ratings (Wetherell et al., 2017; Wetherell & Carter, 2014). When compared to low workload, high workload should elicit a greater increase in autonomic activity and reduce performance, as a result of increased allocation of attentional resources. Based on previous evidence on the functional significance of alpha and theta neural oscillations (Chapter 1.2.4), fronto-medial theta activity should be greatest during high workload, representing engagement of cognitive control processes (Cavanagh & Frank, 2014). Parietal alpha activity should be lower during high workload, representing increased task engagement and allocation of attentional resources (Pfurtscheller, 2006). As manipulations of workload result in modulation of a combination of cognitive resources, no hypotheses regarding the nature of the processes observed can be suggested.

## 2.3 Method

### 2.3.1 Participants

Forty-eight healthy participants (fifteen male, thirty-three female, mean age ± SD = 21.46 ± 3.23 years, range: 18-31 years) took part in this study. Participants were recruited through the University of the West of England's Psychology Participant Pool and local email advertisements. Ethical approval was obtained by the Faculty of Health and Applied Sciences University of the West of England Research Ethics Committee (HAS.18.01.081).

All participants were healthy according to self-reports. Individuals with hypersensitive skin, skin allergies, a pacemaker, and uncorrected vision or hearing were excluded. Due to peripheral and central nervous system measurements, individuals with current or previous anxiety or stress-related disorders, hypertension, neurological, or muscle disorders could not take part. Interested participants were also excluded if they currently took medication apart from over-the-counter analgesia and the contraceptive pill. Smokers were excluded as smoking has been shown to alter autonomic function

(Ashare et al., 2012). Due to further confounding variables on heart rate and perceived stress, there were a number of prerequisites participants were asked to adhere to (Laborde et al., 2017). They were asked to: follow a normal sleep routine the day before the experiment (Stein & Pu, 2012); not partake in intense physical training the day before (Stanley et al., 2013); consume no caffeine, including coffee, chocolate, tea, energy drinks, or headache medication in the two hours before the experiment (Zimmermann-Viehoff et al., 2016); and consume no alcohol in the 24 hours prior to the experiment (Quintana et al., 2013). This self-report data were collected to understand any outliers post-data collection as suggested by Laborde et al. (2017), as these variables may influence HRV parameters.

One person self-declared an endocrine disorder and was excluded from subsequent analyses. Six participants' data were lost due to recording errors. This left 41 participants for the final analysis (twelve male, mean age ± SD = 21.48 ± 3.49 years, range: 18-31 years). Two different participants were electrodermal activity (EDA) non-responders which left 39 participants for subsequent EDA analyses (twelve male, mean age ± SD = 21.48 ± 3.49 years, range: 18-31 years).

### 2.3.2 Stimuli

The Multitasking Framework (Purple Research Solutions, UK) is a cognitively demanding computer task that requires the participant to attend and respond to four tasks simultaneously (see Figure 2.1). Each task is displayed in each quadrant of the monitor. All tasks are performance related and points are awarded for correct responses and deducted for incorrect or missed responses. A running total was presented in the middle of the screen which reflected speed and accuracy of response to the task. 10 points were awarded for every correct answer and 10 points were deducted if the answer was incorrect, or if they did not complete the task in the allocated time. The total score was calculated and recorded each minute. Workload was manipulated by altering the difficulty of the task. In the Multitasking Framework interface, the workload intensity was set as low for the low workload condition, and the workload intensity was set as high for the high workload condition. The four tasks selected were: auditory monitoring, visual monitoring, mental arithmetic, and highest number tap. These tasks were chosen as they require a range of cognitive capacities used in everyday functioning including memory, perception, and attention. These tasks ran simultaneously during a 10-minute duration of each multitasking trial.

Mail alert: this task required participants to identify a high tone and ignore low tones. When a high tone was heard, they were required to click on the 'Incoming Mail' button. In the high workload condition, the time between subsequent tones was approximately five seconds, whereas, during low workload, time between tones was longer at approximately ten seconds. Mental arithmetic: participants were required to add numbers together from right to left. In the high workload condition, participants were required to add together three, three-digit numbers, compared to two, two-digit numbers during low workload. Highest number tap*:* participants identified and selected the highest number within a 4x4 grid of numbers (between 0 and 9). For example, if the highest number in the grid was ‘7’, participants selected all 7s. In the high workload condition, participants had approximately 30 seconds to respond, whereas in the low workload condition, participants had approximately 50 seconds to respond. Bar tracker: this task required participants to identify the height order of bars, once the first bar reached a target line. In the high workload condition, the first bar reached the target line in approximately ten seconds, whereas the first bar reached the target line in approximately 20 seconds during low workload.

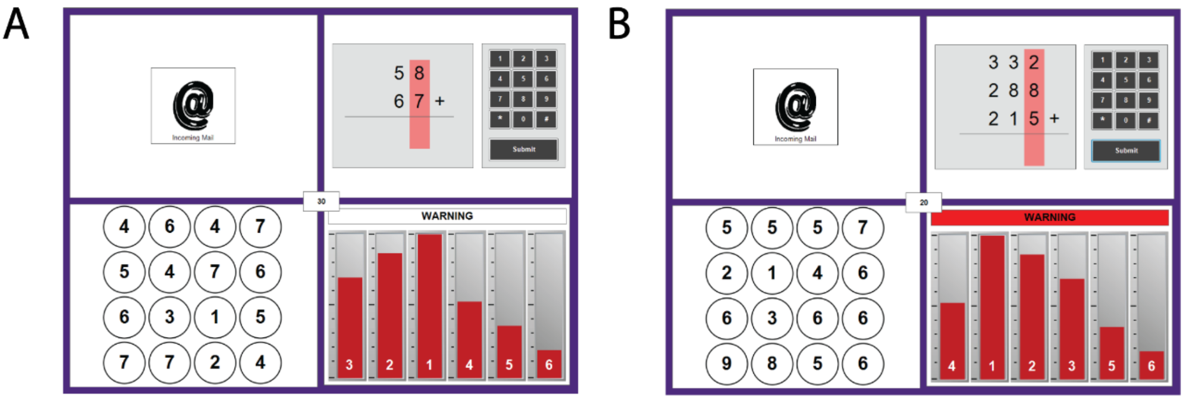



*Figure 2.1.* The Multitasking Framework platform. The multitasking computer task required participants to react to four tasks simultaneously. In both loads, participants partook in mail alert (top left), mental arithmetic (top right), highest number tap (bottom left), bar tracker (bottom right), engaging auditory monitoring, working memory, visual search, and visual monitoring processes respectively. **(A)** Multitasking modules employed during low workload. **(B)** Multitasking modules employed during high workload.

### 2.3.3 Measurements and pre-processing

Several measures were collected including task performance, subjective, physiological (ECG, EMG, EDA), and neurophysiological (EEG). See an overview of measurements collected in Table 2.1 (page 58) and placement of electrodes in Figure 2.2 (page 56) and Figure 2.3 (page 57).

#### *2.3.3.1 Subjective reports (SUB)*

Participants were asked to verbally rate their current “stress” level between 0 and 10 (0 = no stress; 10 = high stress). When the author stated “stress”, the participant verbally responded with a number. Subjective ratings were obtained every one-minute throughout the baseline periods and task. One-minute intervals were chosen as the task mimicked real-world multitasking and as such, it was expected that subjective and physiological measures would adapt to dynamic task demands. Therefore, one-minute epochs were chosen to better capture the evolving relationship between measures and multitasking.

#### *2.3.3.2 Physiological measures*

Electrocardiogram (ECG), electromyogram (EMG), and electrodermal activity (EDA) were recorded by an integrated system and software package (Biopac MP35, BSL Analysis 4.1; Biopac System Inc., California, USA). For ECG and EMG recordings, skin preparation tape was used to lightly abrade the skin and an alcohol pad was used to remove the top layer of dead skin cells. Two disposable self-adhesive with hydrogel Ag/AgCl electrodes (2.5 cm x 2.5 cm; Linton Instrumentation) were placed on the applicable areas (stated below in each subsection). Surgical tape was placed over the electrodes to hold them in place. When the electrodes were applied, the signal quality was checked visually on the screen of the PC. Electrode placement was adjusted if any of the signals were of poor quality.

Pre-processing of ECG, EMG, and EDA data were accomplished using AcqKnowledge 4.4 (BIOPAC Inc., USA). All measurements were averaged over one-minute intervals during the baseline, task, and recovery period. A five-minute duration is considered the gold standard for HRV parameters as the recording should last for 10 times the wavelength of the lower frequency bound (Laborde et al., 2017). However, research has indicated that one-minute averages of time-domain components are accurate and reliable when compared to five-minute recordings, but durations lower than this are less reliable (Esco & Flatt, 2014). Therefore, one-minute epochs were analysed for HRV parameters. To improve the comparability between measures, all physiological measures were averaged over one-minute time periods.

*2.3.3.2.1 Electrocardiogram (ECG)*

The Biopac MP35 amplifier used a band-pass filter of 0.05 Hz and 150 Hz, sampling at 1000 Hz. A standard 2-lead ECG configuration was used with electrodes placed on the right clavicle (negative electrode) and lower left chest (positive electrode). The ECG was checked for signal clarity and recognition of the QRS complex. Data were bandpass filtered with a Finite Impulse Response (FIR) filter between 0.3 Hz and 35 Hz. FIR filters are recommended, as they are more stable compared to Infinite Impulse Response (IIR) filters and are less likely to introduce phase distortions (Cohen, 2014). Next, a peak detection algorithm extracted successive R-waves. All R-waves were visually inspected to determine correct detection. Any peaks incorrectly labelled as an R-wave were removed. Any R-waves missed by the algorithm were labelled manually. If any artefacts or missing data were present, the data were removed from further analyses. Heart rate (HR) values and heart rate variability (HRV) metrics (RMSSD, SDSD, pNN50) were extracted.

*2.3.3.2.2 Electromyogram (EMG)*

The Biopac MP35 amplifier used a band-pass filter of 5 Hz and 250 Hz, sampling at 1000 Hz. Electrodes were placed over the non-dominant trapezius muscle and the spinous process of C7, as described in the recommendations from SENIAM (Hermens et al., 2000). Individual differences in muscle size, thickness or skin can influence EMG amplitude, and so EMG was normalised in relation to a reference maximum voluntary contraction (MVC) value (Roman-Liu, 2016). In line with previous research investigating muscle activity of the trapezius muscle, the participant was instructed to raise their arms 90 degrees, elbows fully extended, wrists straight and palms down for 20 seconds (e.g. Gonçalves et al., 2017; based upon Winkel et al., 1995). Pre-processing of data involved importing data into LabChart (ADInstruments, UK) to remove the ECG components from the EMG signals. ECG-free EMG signals of the left trapezius muscle were subsequently analysed in AcqKnowledge 4.4. Signals were filtered with a FIR bandpass filter of 20 – 450 Hz. A notch filter of 50 Hz and its harmonics were applied. EMG data were visually inspected for artefacts that were indicative of sudden movements (i.e. fast transient spikes). Contaminated data were removed from further analysis. Root-mean-square (RMS-EMG) amplitude values were calculated, with the amplitude normalised to the maximum value of the MVC.

*2.3.3.2.3 Electrodermal activity (EDA)*

The Biopac MP35 amplifier applied a low pass filter of 35 Hz, sampling at 1000 Hz. Participants were first asked to rinse with tap water their non-dominant hand to remove excess oils and dirt. The non-dominant hand was used to reduce the possibility of motion artefacts. Single-point calibration was performed as per the manufacturer's recommendations. Two Ag/AgCl electrodes were mounted in individual polyurethane housings. The electrode cavities were filled with GEL101 isotonic gel (Linton Instrumentation, UK) which acted as an electrolyte for conductance. The electrodes were attached to the distal phalanges of the index and middle finger of the non-dominant hand by a Velcro strap. To ensure the gel had been absorbed, at least five minutes passed before any data was recorded. The participant was asked to breathe in for five seconds and out for five seconds. The participant was considered an EDA non-responder if the EDA signal did not rise after a period of a few seconds. Data were downsampled to 250 Hz for further analysis. A baseline smoothing algorithm was applied in steps of 250 samples/sec. A low pass FIR filter of 5 Hz was applied. Skin conductance responses (SCRs) were quantified as deflections crossing a threshold of 0.03 µs from the background signal. SCRs less than 10% of the maximum peak were discarded from the analysis. Non-specific SCR mean amplitude (NS-SCR amp) and non-specific SCR frequency (NS-SCR freq) values were extracted. As skin conductance level (SCL) values are contaminated by SCRs, the minimum signal SCL value (SCL min) was extracted, as the minimum value will always lie outside of an SCR. This method facilitates a truer representation of tonic changes and has been reported in previous studies (e.g. Braithwaite et al., 2014).

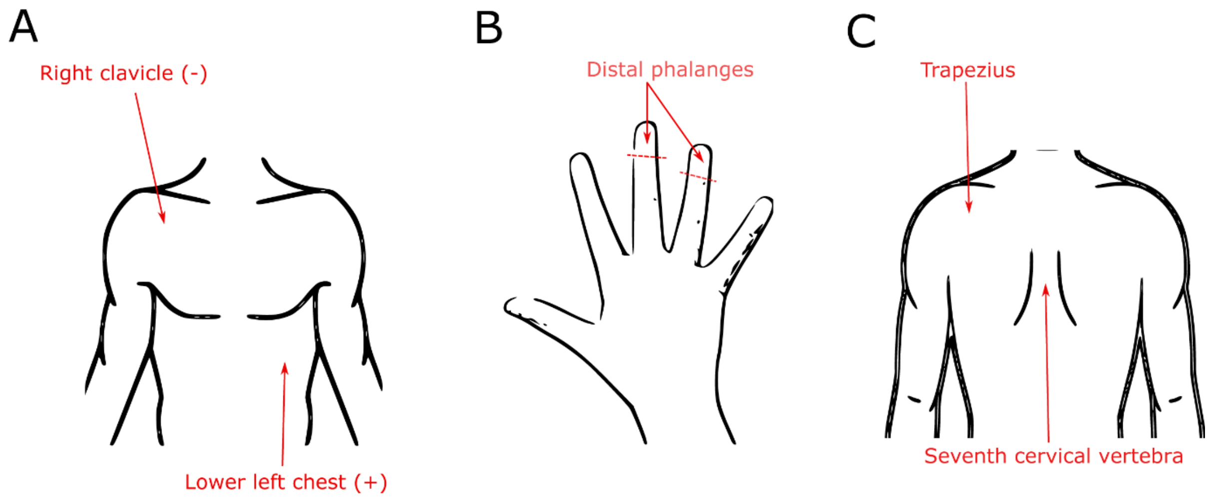



*Figure 2.2.* Placement of electrodes. **(A)** Placement of 2-lead electrocardiogram (ECG). An electrode was placed below the right clavicle and lower left chest, just above and left of the umbilicus. **(B)** Placement of Biopac electrodermal activity (EDA) electrodes. Electrodes were placed on the distal phalanges of the middle and index finger on the non-dominant hand. **(C)** Placement of Biopac electromyogram (EMG) electrodes. An electrode was placed on the non-dominant upper trapezius muscle. A reference electrode was placed over the spinous process of the seventh cervical vertebra.

### *2.3.3.3 Neurophysiological measures*

Electroencephalography (EEG) was recorded using a 128- Channel Geodesic Sensor Net with NetAmps 300 amplifiers (Electrical Geodesics, Eugene, OR). The online sampling rate was set at 1000 Hz and referenced to the vertex (Cz) position. HydroCel saline electrolyte solution, a mixture of potassium chloride with water, was used to improve conductance. The electrical impedance of each electrode was checked and kept below 50 kΩ, either by parting the hair or adding more electrolyte saline solution to the sponges. Participants were shown the raw EEG signals to demonstrate common artefacts that occur due to body and eye movements. Data was first downsampled to 500 Hz and filtered between 0.1 – 100 Hz in NetStation 5, before importing into EEGLAB (Delorme & Makeig, 2004). This was required due to RAM limitations associated with MATLAB. All further pre-processing of EEG data was performed using EEGLAB and custom-written scripts in MATLAB (The MathWorks, Natick, MA, USA). On inspection of the data, mains noise at 50 Hz significantly affected the data and so an FIR filter with a low pass of 48 Hz and high pass of 1 Hz was applied to the data. In dense-array EEG recordings, it is often the case that one or more electrode signals contain prominent artefacts. Therefore, the eeg_interp function was used to replace bad channels with an estimated signal using interpolation from the surrounding electrodes. Signal estimation was carried out using spherical spline interpolation (Perrin et al., 1989). To inform the studies described in Chapters 3 and 4, which utilised a BrainVision 32-channel net, rather than the dense-array 128-channel net used in this study, a 32-channel montage was extracted. Data were then epoched into the necessary segments: baseline low workload, baseline high workload, recovery low workload, recovery high workload (0 – 2 minutes), and task low workload, task high workload (0 – 10 minutes). A marker was added every two seconds, to allow for subsequent epoch segmentation. On each two-second epoch (840 epochs in total), manual trial rejection was conducted to extract major sources of artefacts. Next, Independent Components Analysis (ICA) was computed for further artefact rejection. ICA was computed on each dataset and the topographies and waveforms of the 30 leading components were visualised. One or two leading components that showed a clearly non-cortical origin (mainly vertical and horizontal eye-blinks) were removed. The newly generated data was compared to the previous data to ensure only eye movements were removed and the neural signal was left. Once artefacts were removed, the data were re-referenced to the average of all EEG channels. The data were then bandpass filtered using an FIR filter into theta (4 – 7 Hz) and alpha (8.5 – 12.5 Hz). To assess changes in oscillatory power, the Hilbert transform was applied to compute the envelope of the amplitude-modulated signal. Therefore, phase information was discarded. Next, the data were converted into their absolute values. Data were normalised by taking the natural logarithm of spectral power. Since previous studies have focused on frontal and parietal areas when comparing alpha and theta activity related to attentional processing (see Chapter 1.2.4), two electrode clusters were extracted based on these areas. For the frontal cluster, data from electrodes F3, F4, F7, F8, and Fz were extracted and averaged together. For the parietal cluster, data from electrodes P3, P4, P7, P8, and Pz were extracted and averaged together. These specific electrodes were chosen based upon previous research utilising similar multitasking paradigms (e.g. Puma et al., 2018; see Figure 2.3).

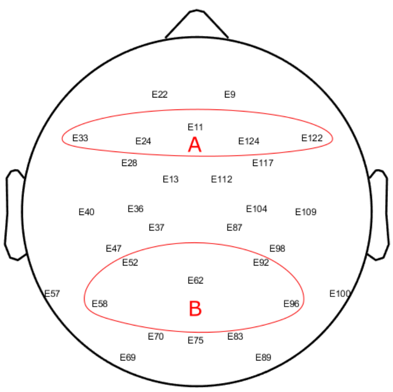



*Figure 2.3.* Extracted electrodes of interest. **(A)** Frontal cluster. **(B)** Parietal cluster.

*Table 2.1.* A summary of the measures collected. Electrophysiological recordings were collected continuously. Verbal ratings of perceived stress were collected every one-minute during baseline, task period, and recovery. Multitasking Framework score was collected every minute during multitasking and resulted in a total score at the end of each workloadassessment.

| Measurement | Parameter | Abbreviation |
|---|---|---|
| Electrocardiography | Heart rate | HR |
| | Root-mean-square of successive RR differences | RMSDD |
| | The percentage of adjacent normal RR intervals that differ from each other by more than 50ms | pNN50 |
| | Standard deviation of normal RR intervals | SDSD |
| Electrodermal activity | Minimum skin conductance level amplitude | SCL min |
| | Non-specific skin conductance response amplitude | NS-SCR amp |
| | Non-specific skin conductance response frequency | NS-SCR freq |
| Electromyography | Root-mean-square | RMS-EMG |
| Electroencephalography | Frontal alpha power | - |
| | Parietal alpha power | - |
| | Frontal theta power | - |
| | Parietal theta power | - |
| Verbal rating | Stress rating | SUB |
| Multitasking Framework | Cumulated score | - |
| | Total score | - |

### 2.3.4 Procedure

In line with previous studies utilising the Multitasking Framework (e.g. Wetherell et al., 2017), all data collection took place between 12.00 and 16:00 due to the circadian rhythm of heart rate and cortisol (Dickmeis, 2009; van Eekelen et al., 2004). Following initial screening at enrolment to check that participants met the inclusion criteria, participants arrived at the laboratory and filled in a demographic questionnaire. Participants were given a two-minute demonstration of the multitasking framework to make sure they understood how to perform the tasks. The electrophysiological recording equipment was set up, including the EMG, EDA, ECG and EEG electrodes. The physiological data were recorded continuously from set up to the end of the last condition.

A repeated measures protocol required participants to partake in four conditions: standing, sitting, low workload, and high workload (see Figure 2.4). The sitting and standing data are not commented on further in this thesis. Conditions were performed in a counterbalanced order so that each possible permutation appeared twice. The standing and sitting condition lasted five minutes each. The low and high workload condition lasted ten minutes each preceded by a two-minute baseline measurement, followed by a two-minute recovery measurement. During the baseline and recovery measurement periods, participants were asked to relax and keep their eyes open. Perceived stress levels were assessed every minute in the low workload and high workload conditions.

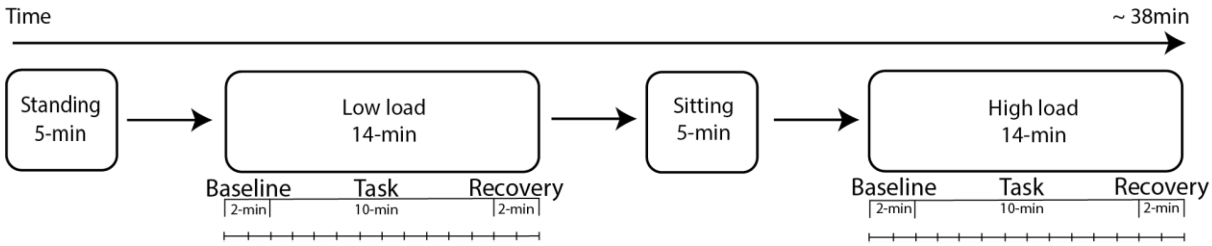



*Figure 2.4.* The experimental procedure. The order of conditions (standing, sitting, low workload, high workload) were counterbalanced across participants. Physiological data were subsequently split up into one-minute epochs, equalling 14 epochs each in the low and high workload conditions.

### 2.3.5 Statistical analyses

Shapiro-Wilk test of normality and visualisation of QQ plots of the unstandardized residuals indicated that the subjective, HR, HRV, and SCR data were not normally distributed. Therefore, data were normalised to make them suitable for parametric analyses. All data were initially normalised via the natural logarithmic transformation. The normalisation correction added a constant (1) if the value could involve a zero response (subjective ratings, pNN50, NS-SCR freq, NS-SCR amp). This method of normalisation is reported in previous studies (e.g. Stanley et al., 2013), and is recommended for heart rate, heart rate variability, and skin conductance measures (Boucsein et al., 2012; Laborde et al., 2017).

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 25 (IBM Corp., Armonk, N.Y., USA). A paired-samples *t*-test was undertaken on multitasking total score performance data. Subjective, psychophysiological, and EEG data were submitted to a mixed model analysis with the maximum likelihood method (Hoffman & Rovine, 2007). A linear mixed-effects model was used due to the uneven number of epochs to be compared (i.e. two one-minute epochs for baseline and recovery periods, and ten one-minute epochs for the task period), which can be handled by mixed models (Singmann & Kellen, 2019). Fixed effects were Workload (low, high), Interval (baseline, task, recovery), and Time ([1 2], [1 2 3 4 5 6 7 8 9 10], [1 2]). The maximum likelihood method was chosen to compare models and interpret the fixed effects (Field, 2013, p. 835). A sequence of models were fit with gradually decreasing complexity of the random effect structure (Baayen et al., 2008). A random intercept and slope model consisting of Time as a random effect was compared against a random intercept model at the participant level. The −2 log likelihood of different models was calculated in order to determine the optimal statistical complexity model. In all cases, a random intercept and slope model provided a better fit for the data. This model assumes that the slope across Time varies across participants. To examine which covariance structure yielded the best fit for the dataset, various covariance structures were tested: variance components, diagonal, autoregressive structure, and unstructured. In all cases except RMS-EMG, an autoregressive structure yielded the best fit. For RMS-EMG, a diagonal covariance structure was applied. As the interaction between the fixed effects were of interest, the *F* tests are subsequently reported, with *post hoc* analyses run with Bonferroni correction.

The statistical threshold for significance was set to $p < 0.05$. Effect sizes were reported as Cohen's $f^2$ for fixed effects, as advised by Lorah (2018) and Selya et al. (2012). An effect is considered small at a value of 0.02, medium at a value of 0.15, and large at a value of 0.35 (Cohen, 1992). For paired-samples *t*-tests, Cohen's $d_z$ was reported. Although no benchmarks have been set, Lakens (2017) previously suggested 0.14 for a small effect, 0.35 for a medium effect, and 0.57 for a large effect.

## 2.4 Results

First, task performance was compared between low workload and high workload. Next, a random intercept and slope mixed model was undertaken on subjective, psychophysiological, and EEG data, with fixed effects: Workload (low, high), Interval (baseline, task, recovery), and Time ([1 2], [1 2 3 4 5 6 7 8 9 10], [1 2]). Table 2.2 summaries natural logarithm transformed means (SD) of subjective and physiological data at one-minute intervals during baseline, task, and recovery during low and high workload. Appendix 2.1 reports the raw values for ease of interpretation. Figure 2.6 provides a graphical representation of subjective and psychophysiological measures at one-minute intervals during baseline, task, and recovery intervals during low and high workload multitasking. For EEG data, Table 2.3 summarises the means (SD) over interval periods during low and high workload multitasking, and Appendix 2.2 provides the raw values for ease of interpretation. Appendix 2.3 provides a more comprehensive summary of the means (SD) at one-minute intervals during baseline, task, and recovery during low and high workload, while Appendix 2.4 reports the raw values. Figure 2.7 provides a graphical representation of alpha and theta results.

### 2.4.1 Task performance

A paired-samples *t*-test revealed a significant difference in task performance between the low workload and high workload condition, $t_{(40)} = 10.51$, $p < .001$, $d_z = 1.64$. Participants scored higher during low workload ($M = 1075.12$, $SD = 229.47$) compared to high workload ($M = 740.49$, $SD = 234.45$; see Figure 2.5).

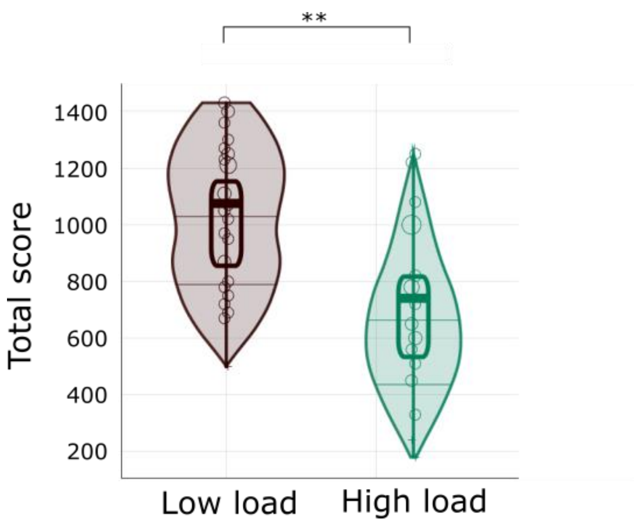



*Figure 2.5.* Violin plot representing total score during the Multitasking Framework. The black bolded line represents the mean. The box surrounding the mean represents the interquartile range, and the vertical lines represent the lower/upper adjacent values. On each side of the box is a kernel density estimation that shows the distribution of the data. The wider the section, the higher the probability that members of the population will take on the given value. In the low workload condition, participants scored higher, but were more variable in performance. Key: ** $p < .001$

### 2.4.2 Subjective reports (SUB)

The mixed-effects model revealed a significant main effect for Workload, $F_{(1, 984.23)} = 95.32$, $p < .001$, $f^2 = 0.31$, Interval, $F_{(2, 984.23)} = 282.58$, $p < .001$, $f^2 = 0.76$, and Time, $F_{(9, 648.48)} = 6.70$, $p < .001$, $f^2 = 0.28$ on subjective ratings of stress. In addition, a statistically significant interaction between Workload and

Interval, $F_{(2,\ 984.23)} = 438.10$, $p < .001$, $f^2 = 0.94$, and Interval and Time, $F_{(1,\ 984.23)} = 21.73$, $p < .001$, $f^2 = 0.15$, were revealed. The interaction between Workload and Time, $F_{(1,\ 984.23)} = 0.34$, $p = .96$, and the three-way interaction, $F_{(2,\ 984.23)} = 0.80$, $p = .45$ were not significant. Pairwise comparisons of the Workload by Interval interaction revealed SUB during baseline did not differ between low and high workload ($p = .12$). SUB ratings were significantly greater during the high workload task ($p < .001$) and recovery ($p < .001$), when compared to low workload. Interval by Time *post hoc* comparisons revealed that during the task, SUB ratings were greater during the first minute than the second minute ($p < .001$). Minutes three to five exhibited greater ratings than the first minute ($p = .001$). Minutes six to ten revealed greater ratings compared to the first and second minute ($p = .007$ to $p < .001$). Overall, these findings demonstrate that verbal stress ratings were greater during the task and at recovery in the high workload condition, when compared to low workload. During the task, verbal stress ratings were high in the first minute, decreased after the second minute, but then increased again throughout the task.

### 2.4.3 Heart rate (HR)

The mixed-effects model revealed a simple main effect for Time on HR, $F_{(9,\ 677.04)} = 5.10$, $p < .001$, $f^2 = 0.23$. The main effects for Workload, $F_{(1,\ 1028.62)} = 2.95$, $p = .09$, and Interval, $F_{(2,\ 1028)} = 0.68$, $p = .51$ were not significant. Similarly, neither were the interaction effects between Workload and Interval, $F_{(2,\ 1027.7)} = 1.92$, $p = .15$, Workload and Time, $F_{(9,\ 1028.45)} = 0.46$, $p = .91$, Time and Interval, $F_{(2,\ 1027.70)} = 2.01$, $p = .13$, or the three-way interaction, $F_{(2,\ 1027.7)} = 0.19$, $p = .83$. *Post hoc* comparisons revealed HR was greater at minute one than minute two ($p < .001$). HR was statistically greater at minutes five to ten, compared to time point two ($p = .05$ to $p < .001$). In summary, *post hoc* comparisons revealed that after the fifth minute of the task, heart rate statistically increased, reflecting increased arousal, until the end of the task.

### 2.4.4 Heart rate variability: RMSSD

The mixed-effects model revealed a statistically significant effect for Workload on RMSSD, $F_{(1,\ 1030.47)} = 4.95$, $p = .03$, $f^2 = 0.06$. The main effects for Interval, $F_{(2,\ 1029.47)} = 2.64$, $p = .07$, and Time, $F_{(9,\ 716.03)} = 0.74$, $p = .67$ were not significant. The interactions between Workload and Interval, $F_{(2,\ 1029.47)} = 1.30$, $p = .27$, Workload and Time, $F_{(9,\ 1030.21)} = 0.27$, $p = .98$, Interval and Time, $F_{(2,\ 10.29.47)} = 0.67$, $p = .51$, and the three-way interaction, $F_{(2,\ 1029.47)} = 0.36$, $p = .97$, were also not significant. *Post hoc* comparisons revealed that RMSSD was greater during high compared to low workload ($p = .02$). These results suggest that RMSSD was greater during the high workload condition, reflecting increased variability in heart rate, irrespective of Time or Interval.

### 2.4.5 Heart rate variability: SDSD

The mixed-effects model revealed a statistically significant effect for Workload on SDSD, $F_{(1, 1030.51)} = 4.94$, $p = .03$, $f^2 = 0.06$. No other effects were statistically significant: Interval, $F_{(2, 1029.52)} = 2.64$, $p = .07$, Time, $F_{(9, 716.15)} = 0.74$, $p = .68$, Workload and Time, $F_{(9, 1030.25)} = 0.27$, $p = .98$, Interval and Time, $F_{(2, 1029.52)} = 0.67$, $p = .51$, Workload, Interval and Time, $F_{(2, 1029.52)} = 0.04$, $p = .97$. *Post hoc* comparisons revealed that SDSD was greater during high compared to low workload ($p = .02$). These results suggest that SDSD was greater during the high workload condition, reflecting increased variability in heart rate, irrespective of Time or Interval.

### 2.4.6 Heart rate variability: pNN50

Similarly to the other heart rate variability measures, the mixed-effects model revealed a statistically significant effect for Workload on pNN50, $F_{(1, 997.56)} = 10.12$, $p = .002$, $f^2 = 0.10$. Again, no other effect yielded statistical significance: Interval, $F_{(2, 967.56)} = 0.17$, $p = .84$, Time, $F_{(9, 608.8)} = 1.25$, $p = .26$, Workload and Time, $F_{(9, 997.56)} = 1.31$, $p = .27$, Interval and Time, $F_{(2, 997.56)} = 0.44$, $p = .64$), Workload, Interval and Time, $F_{(2, 997.56)} = 0.14$, $p = .89$. *Post hoc* comparisons revealed that pNN50 was greater during high compared to low workload ($p = .005$). These results suggest that pNN50 was greater during the high workload condition, reflecting increased variability in heart rate, irrespective of Time or Interval.

### 2.4.7 Skin conductance response frequency (NS-SCR freq)

A statistically significant main effect for Interval, $F_{(2, 937.38)} = 18.07$, $p < .001$, $f^2 = 0.07$, and Time, $F_{(9, 633.63)} = 11.4$, $p < .001$, $f^2 = 0.38$ on NS-SCR freq was found. The model also revealed an interaction effect between Workload and Interval, $F_{(2, 937.38)} = 3.58$, $p = .03$, $f^2 = 0.07$. The main effect of Workload was not significant, $F_{(1, 937.38)} = 0.03$, $p = .87$. The interaction effects between Workload and Time, $F_{(9, 937.38)} = 0.54$, $p = .85$, Interval and Time, $F_{(2, 937.38)} = 1.35$, $p = .26$, and Workload, Interval and Time, $F_{(2, 937.38)} = 0.15$, $p = .87$ were also not significant. For the interaction effect, *post hoc* comparisons revealed that during high workload, there were no differences in NS-SCR freq over intervals. However, during low workload, NS-SCR freq was greater in the baseline period, compared to the task ($p < .001$). This suggests that there were more SCRs, reflecting increased arousal, before task commencement, irrespective of workload.

### 2.4.8 Skin conductance response amplitude (NS-SCR amp)

A significant main effect for Time on NS-SCR amp was found, $F_{(9, 601.82)} = 6.20$, $p \leq .001$, $f^2 = 0.28$. No other significant effects were found: Workload, $F(1, 927.62) = 3.43$, $p = .06$, Interval, $F_{(2, 927.62)} = 2.20$, $p = .11$, Workload and Interval, $F_{(2, 927.62)} = 0.62$, $p = .54$, Workload and Time, $F_{(9, 927.62)} = 0.71$, $p = .70$, Interval and Time, $F_{(2, 927.62)} = 1.45$, $p = .23$, Workload, Interval and Time, $F_{(2, 927.62)} = 0.36$, $p = .70$. Pairwise comparisons revealed that NS-SCR amp was greater during the first minute, when compared to all other minutes apart from the seventh minute ($p \leq .001$ to $p = .003$). Therefore, regardless of load, NS-SCR amp was greater during the first minute compared to all other time points besides the seventh minute, reflecting increased arousal.

### 2.4.9 Skin conductance level (SCL min)

The model revealed a significant main effect for Interval, $F_{(2, 960.44)} = 18.23$, $p < .001$, $f^2 = 0.19$, and Time, $F_{(9, 960.44)} = 617.55$, $p < .001$, $f^2 = 2.39$, on SCL min. The mixed-effects model revealed a significant interaction between Interval and Time, $F_{(2, 960.44)} = 3.62$, $p = .03$, $f^2 = 0.07$. No other effects were significant: Workload, $F_{(1, 960.44)} = 0.63$, $p = .43$, Workload and Interval, $F_{(2, 960.44)} = 0.9$, $p = .41$, Workload and Time, $F_{(9, 960.44)} = 0.06$, $p = 1.0$; Workload, Interval and Time, $F_{(2, 960.44)} = 0.02$, $p = .98$. *Post hoc* comparisons uncovering the interaction effect revealed that during the task, SCL at the first minute was greater than minutes seven to ten ($p = .01$ to $p < .001$). SCL during the second minute was greater compared to minutes four to ten ($p = .003$ to $p < .001$). The third minute displayed higher SCL values compared to minutes seven to ten ($p = .003$ to $p = .05$). After the task had ended, during the recovery period, the second minute displayed larger SCL values compared to the first minute ($p = .04$). These findings suggest that during the task, SCL rose rapidly, reflecting heightened arousal, during the first and second minute, regardless of load. Once the task finished, SCL was greater during the second compared to the first minute of rest.

### 2.4.10 Root-mean-square of electromyography (RMS-EMG)

A significant main effect for Workload, $F_{(1, 1005.9)} = 12.6$, $p \leq .001$, $f^2 = 0.12$, Interval, $F_{(2, 1005.9)} = 48.74$, $p \leq .001$, $f^2 = 0.11$, and Time, $F_{(9, 230.47)} = 2.52$, $p = .009$, $f^2 = 0.05$, on RMS-EMG was revealed. The interactions Workload and Interval, $F_{(2, 1005.9)} = 2.15$, $p = .12$, Workload and Time, $F_{(9, 1005.9)} = 0.77$, $p = .65$, Interval and Time, $F_{(2, 1005.9)} = 2.72$, $p = .07$, and the three-way interaction, $F_{(2, 1005.9)} = 0.35$, $p = .70$, were not significant. Pairwise comparisons revealed that high workload displayed greater RMS-EMG values compared to low workload ($p < .001$). When compared to the active task period, RMS-EMG values were lower during baseline ($p < .001$), and during recovery ($p < .001$). The main effect of Time revealed that EMG activity was greater during minutes three to ten when compared to the first minute ($p < .001$ to $p = .02$) and the second minute ($p < .001$). These results reveal that RMS-EMG values increased throughout the task and were greater during high workload.

*Table 2.2.* Table summarising natural logarithm transformed means (SD) of subjective and physiological data at one-minute intervals during baseline, task, and recovery during low and high workload.

| Measure | Workload | Minute periods over task | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Baseline | | Task | | | | | | | | | | Recovery | |
| | | 1 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 1 | 2 |
| Subjective (ln) | Low | 0.82 | 0.82 | 1.30 | 1.40 | 1.50 | 1.53 | 1.52 | 1.57 | 1.63 | 1.61 | 1.60 | 1.65 | 1.09 | 0.90 |
| | | (0.49) | (0.49) | (0.40) | (0.41) | (0.42) | (0.43) | (0.53) | (0.42) | (0.42) | (0.46) | (0.46) | (0.44) | (0.59) | (0.51) |
| | High | 0.90 | 0.86 | 1.56 | 1.69 | 1.80 | 1.82 | 1.85 | 1.87 | 1.87 | 1.88 | 1.87 | 1.88 | 1.30 | 1.00 |
| | | (0.59) | (0.60) | (0.59) | (0.40) | (0.32) | (0.35) | (0.35) | (0.36) | (0.37) | (0.35) | (0.38) | (0.37) | (0.52) | (0.52) |
| Heart rate (ln) | Low | 4.30 | 4.28 | 4.30 | 4.28 | 4.29 | 4.30 | 4.30 | 4.30 | 4.30 | 4.31 | 4.32 | 4.31 | 4.31 | 4.28 |
| | | (0.13) | (0.13) | (0.13) | (0.13) | (0.13) | (0.13) | (0.13) | (0.13) | (0.13) | (0.12) | (0.12) | (0.13) | (0.13) | (0.13) |
| | High | 4.29 | 4.28 | 4.30 | 4.29 | 4.29 | 4.30 | 4.30 | 4.31 | 4.30 | 4.31 | 4.30 | 4.31 | 4.31 | 4.28 |
| | | (0.14) | (0.14) | (0.13) | (0.13) | (0.13) | (0.13) | (0.12) | (0.11) | (0.12) | (0.12) | (0.12) | (0.12) | (0.13) | (0.13) |
| RMSSD (ln) | Low | 4.01 | 3.89 | 3.86 | 3.84 | 3.80 | 3.74 | 3.82 | 3.82 | 3.93 | 3.78 | 3.85 | 3.77 | 3.96 | 3.81 |
| | | (0.78) | (0.85) | (0.66) | (0.66) | (0.70) | (0.66) | (0.83) | (0.90) | (1.00) | (0.66) | (0.71) | (0.75) | (0.79) | (0.63) |
| | High | 4.04 | 3.94 | 3.89 | 3.89 | 3.90 | 3.87 | 3.83 | 3.83 | 3.88 | 3.85 | 3.87 | 3.86 | 4.11 | 4.02 |
| | | (0.91) | (0.86) | (0.82) | (0.79) | (0.83) | (0.92) | (0.79) | (0.81) | (0.77) | (0.77) | (0.83) | (0.80) | (0.82) | (0.81) |
| | | 4.01 | 3.89 | 3.86 | 3.84 | 3.80 | 3.74 | 3.82 | 3.82 | 3.93 | 3.78 | 3.85 | 3.77 | 3.96 | 3.81 |
| SDSD (ln) | Low | (0.79) | (0.86) | (0.67) | (0.67) | (0.84) | (0.91) | (1.01) | (0.67) | (0.72) | (0.76) | (0.72) | (0.73) | (0.80) | (0.64) |
| | High | 4.04 | 3.94 | 3.89 | 3.89 | 3.90 | 3.87 | 3.84 | 3.83 | 3.88 | 3.85 | 3.87 | 3.86 | 4.10 | 4.02 |
| | | (0.64) | (0.92) | (0.87) | (0.83) | (0.80) | (0.84) | (0.93) | (0.80) | (0.82) | (0.78) | (0.84) | (0.81) | (0.83) | (0.82) |
| pNN50 (ln) | Low | 2.96 | 2.84 | 2.95 | 2.96 | 2.82 | 2.71 | 2.72 | 2.83 | 2.73 | 2.80 | 2.70 | 2.71 | 2.90 | 2.90 |
| | | (1.13) | (1.29) | (1.05) | (1.21) | (1.30) | (1.36) | (1.43) | (1.25) | (1.36) | (1.27) | (1.34) | (1.34) | (1.08) | (1.12) |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | High | 3.06 | 2.98 | 2.92 | 2.93 | 2.86 | 2.76 | 2.82 | 2.79 | 2.97 | 2.85 | 2.80 | 2.95 | 3.10 | 3.00 |
| | | (1.24) | (1.30) | (1.17) | (1.21) | (1.33) | (1.34) | (1.28) | (1.35) | (1.16) | (1.24) | (1.37) | (1.15) | (1.09) | (1.19) |
| NS-SCR freq (ln) | Low | 1.19 | 1.02 | 1.37 | 1.08 | 0.78 | 0.86 | 0.87 | 0.85 | 0.85 | 0.76 | 0.83 | 0.81 | 1.15 | 0.90 |
| | | (0.59) | (0.53) | (0.50) | (0.68) | (0.70) | (0.65) | (0.63) | (0.66) | (0.57) | (0.67) | (0.67) | (0.69) | (0.45) | (0.56) |
| | High | 1.01 | 0.88 | 1.48 | 1.16 | 1.07 | 0.97 | 0.94 | 0.91 | 0.96 | 0.89 | 0.88 | 0.87 | 1.08 | 0.92 |
| | | (0.59) | (0.58) | (0.56) | (0.65) | (.59) | (0.66) | (0.61) | (0.52) | (0.5) | (0.59) | (0.54) | (0.57) | (0.48) | (0.56) |
| NS-SCR amp (ln) | Low | 2.00 | 1.92 | 2.25 | 1.81 | 1.45 | 1.56 | 1.63 | 1.60 | 1.68 | 1.40 | 1.49 | 1.43 | 2.09 | 1.78 |
| | | (0.81) | (0.82) | (0.50) | (0.95) | (1.11) | (1.02) | (1.00) | (1.04) | (1.00) | (1.08) | (1.03) | (1.09) | (1.06) | (0.94) |
| | High | 1.90 | 1.75 | 2.18 | 1.92 | 1.85 | 1.68 | 1.77 | 1.79 | 1.89 | 1.68 | 1.73 | 1.67 | 2.05 | 1.80 |
| | | (0.92) | (1.00) | (0.72) | (0.94) | (0.89) | (1.00) | (0.93) | (0.87) | (0.77) | (0.96) | (0.92) | (0.95) | (0.70) | (0.96) |
| SCL min (µS) | Low | 7.21 | 7.01 | 7.41 | 7.78 | 7.41 | 7.05 | 6.92 | 6.89 | 6.8 | 6.71 | 6.65 | 6.58 | 6.71 | 7.10 |
| | | (3.04) | (3.05) | (3.12) | (2.84) | (2.82) | (2.70) | (2.74) | (3.01) | (2.88) | (2.89) | (2.81) | (2.98) | (2.00) | (2.78) |
| | High | 7.15 | 6.91 | 7.61 | 7.82 | 7.44 | 7.14 | 7.02 | 6.96 | 6.84 | 6.81 | 6.70 | 6.70 | 6.98 | 7.28 |
| | | (3.08) | (2.97) | (3.26) | (2.76) | (2.57) | (2.50) | (2.66) | (2.60) | (2.52) | (2.57) | (2.54) | (2.64) | (2.89) | (2.97) |
| RMS-EMG (%) | Low | 25.7 | 23.05 | 32.89 | 33.00 | 33.45 | 34.76 | 35.83 | 36.62 | 37.18 | 36.52 | 38.17 | 39.99 | 27.26 | 24.32 |
| | | (13.99) | (14.16) | (15.92) | (17.36) | (17.64) | (18.82) | (20.73) | (22.49) | (21.89) | (20.42) | (21.39) | (21.83) | (16.91) | (17.45) |
| | High | 27.21 | 25.19 | 35.28 | 38.5 | 39.70 | 40.70 | 41.08 | 40.86 | 41.45 | 45.20 | 42.46 | 41.37 | 26.86 | 23.36 |
| | | (15.39) | (13.69) | (18.01) | (20.86) | (20.33) | (20.85) | (20.99) | (20.63) | (21.22) | (28.99) | (24.11) | (23.94) | (17.49) | (16.43) |

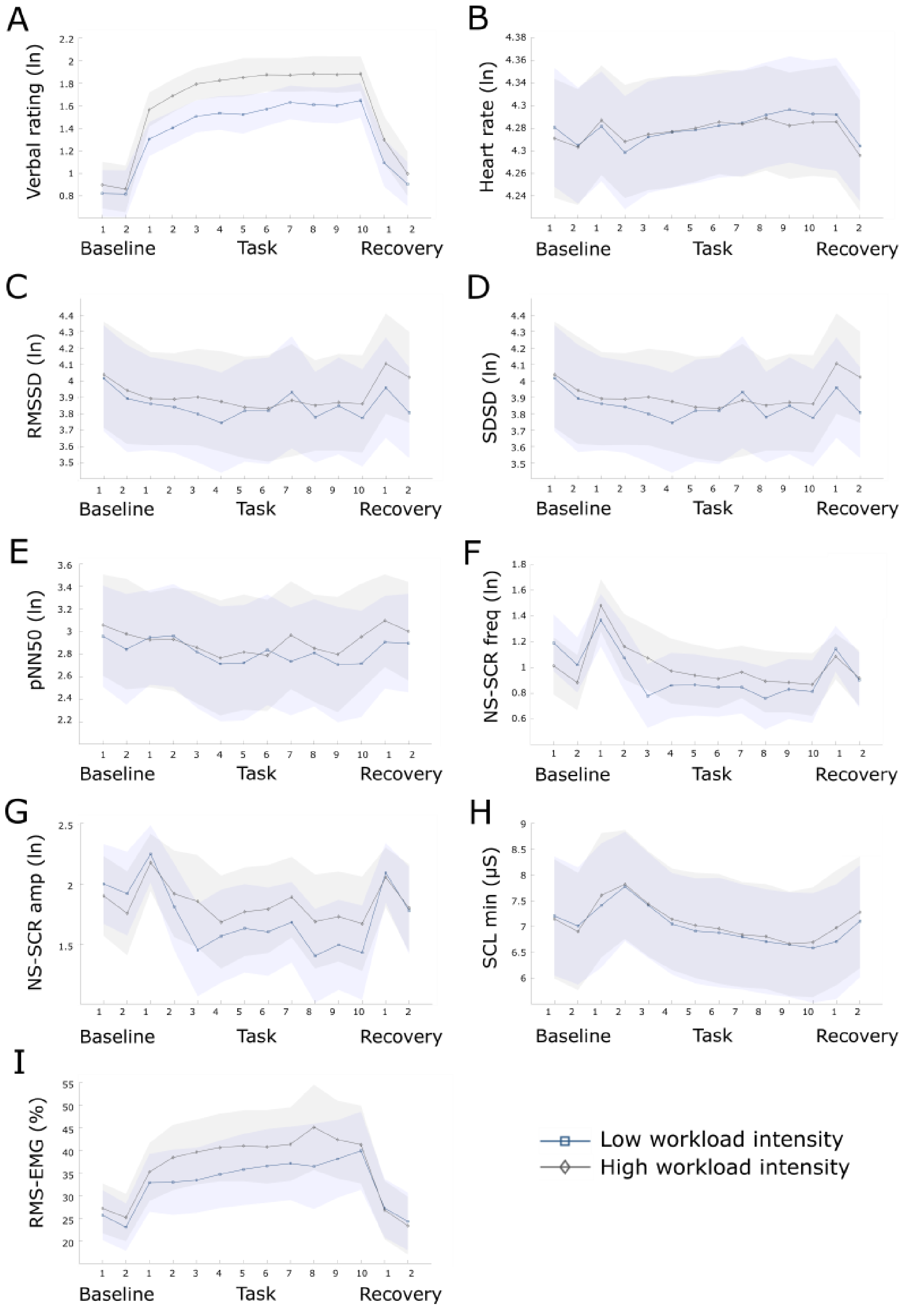



*Figure 2.6.* Subjective and psychophysiological measures at one-minute intervals during baseline, task, and recovery intervals during low and high workload multitasking. Shaded areas represent the ± standard error of the mean difference. **(A)** Natural logarithm transformed subjective ratings. **(B)** Natural logarithm transformed averaged heart rate. **(C)** Natural logarithm transformed averaged root-mean-square of successive differences

between normal heartbeats (RMSSD). **(D)** Natural logarithm transformed averaged standard deviation of the successive differences between normal heartbeats (SDSD). **(E)** Natural logarithm transformed averaged proportion of the number of pairs of normal heartbeats that differ by more than 50ms divided by total number of heartbeats (pNN50). **(F)** Natural logarithm transformed averaged frequency of non-specific skin conductance responses (NS-SCR freq). **(G)** Natural logarithm transformed averaged amplitude of non-specific skin conductance responses (NS-SCR amp). **(H)** Average minimum skin conductance level (µS; SCL min). **(I)** Averaged normalised root-mean-square of the electromyogram signal (RMS-EMG).

### 2.4.11 Alpha oscillatory activity

A linear mixed model was run to understand the impact of workload on frontal alpha activity. The model revealed a significant main effect for Workload, $F_{(1, 867.85)} = 34.98$, $p < .001$, $f^2 = 0.20$, Interval, $F_{(2, 867.85)} = 45.66$, $p < .001$, $f^2 = 0.32$, and Time, $F_{(9, 438.51)} = 2.18$, $p = .03$, $f^2 = 0.15$. There were no significant interaction effects: Workload and Interval, $F_{(2, 867.85)} = 0.37$, $p = .70$, Workload and Time, $F_{(9, 867.85)} = 0.06$, $p = 1.0$, Interval and Time, $F_{(2, 867.85)} = 0.19$, $p = .83$, or the three-way interaction, $F_{(2, 867.85)} = .25$, $p = .78$. *Post hoc* comparisons revealed that alpha activity was greater during low compared to high workload ($p < .001$). Comparisons for Interval revealed alpha was lower at baseline compared to recovery ($p < .001$). In addition, alpha activity was lower during the task compared to recovery ($p < .001$). There were no significant differences between time points once multiple comparisons were corrected for.

A second linear mixed model was run to understand the impact of workload on parietal alpha activity. The mixed model revealed a significant main effect for Workload, $F_{(1, 869.89)} = 31.18$, $p < .001$, $f^2 = 0.19$, Interval, $F_{(2, 869.89)} = 50.47$, $p \leq .001$, $f^2 = 0.34$, and Time, $F_{(9, 349.1)} = 2.84$, $p = .003$, $f^2 = 0.22$. Similarly to frontal activity, there were no interaction effects: Workload and Interval, $F_{(2, 869.89)} = 0.20$, $p = .82$, Workload and Time, $F_{(9, 869.89)} = 0.12$, $p = 1.0$, Interval and Time, $F_{(2, 869.89)} = 0.35$, $p = .71$, Workload, Interval and Time, $F_{(2, 869.89)} = 0.18$, $p = .84$. *Post hoc* comparisons revealed that alpha activity was greater during low workload compared to high ($p < .001$). Comparisons for Interval revealed alpha was lower at baseline compared to recovery ($p < .001$). In addition, alpha activity was lower during the task compared to recovery ($p < .001$). There was no significant difference between time points once multiple comparisons were corrected for.

Overall, frontal and parietal alpha were highest during low workload, regardless of interval. During the task, frontal and parietal alpha activity did not differ from baseline levels. Both frontal and parietal alpha increased following task cessation during the recovery period.

2.4.12 Theta oscillatory activity

A linear mixed model was run to understand the impact of workload on frontal theta activity. The model revealed a significant main effect for Workload, $F_{(1, 890.71)} = 41.52$, $p < .001$, $f^2 = 0.21$, a significant main effect for Interval, $F_{(2, 890.71)} = 10.89$, $p < .001$, $f^2 = 0.15$, and a significant main effect for Time, $F_{(9, 878.66)} = 2.45$, $p = .009$, $f^2 = 0.12$. No interaction effects were significant: Workload and Interval, $F_{(2, 890.71)} = 0.50$, $p = .61$, Workload and Time, $F_{(9, 890.71)} = 0.04$, $p = 1.0$, Interval and Time, $F_{(2, 890.71)} = 0.39$, $p = .67$, or Workload, Interval and Time, $F_{(2, 890.71)} = 0.41$, $p = .67$. *Post hoc* comparisons revealed that frontal theta was significantly greater during low workload, compared to high workload ($p < .001$). Comparisons for Interval revealed that, during the task, theta was significantly greater than during baseline ($p < .001$). During recovery, theta was significantly greater than baseline ($p < .001$). Time effects revealed that theta was greater at minute ten when compared to minute one ($p < .001$) and minute two ($p = .004$); and greater at minute nine when compared to minute one ($p = .03$).

A final linear mixed model was run to understand the impact of workload on parietal theta activity. The model revealed a significant main effect for Workload, $F_{(1, 917.97)} = 32.88$, $p < .001$, $f^2 = 0.19$, a significant main effect for Interval, $F_{(2, 917.97)} = 20.39$, $p < .001$, $f^2 = 0.21$, and a significant main effect for Time, $F_{(9, 916.81)} = 3.82$, $p < .001$, $f^2 = 0.17$. No interaction effects were significant: Workload and Interval, $F_{(2, 917.97)} = .43$, $p = .65$, Workload and Time, $F_{(9, 917.97)} = 0.16$, $p = 1.0$, Interval and Time, $F_{(2, 917.97)} = 0.45$, $p = .64$, Workload, Interval and Time, $F_{(2, 917.97)} = 0.32$, $p = .73$. *Post hoc* comparisons revealed that parietal theta was significantly greater during low workload compared to high workload ($p < .001$). Comparisons for Interval revealed that during the task, theta was significantly greater than baseline ($p < .001$). Theta was significantly greater during recovery when compared to baseline ($p < .001$). Time effects revealed that theta was greater at minute nine when compared to minute one ($p < .02$) and two ($p < .03$). Theta was also greater at minute ten when compared to minute one ($p < .001$) and two ($p < .001$).

In summary, frontal and parietal theta revealed highest power values during low workload, regardless of interval. During the task, frontal and parietal theta activity increased, and stayed at a similar level following task cessation, during the recovery period.

*Table 2.3*. Table summarising natural logarithm transformed means (SD) of alpha and theta activity averaged during baseline, task, and recovery of low and high workload multitasking.

| Frequency (Hz) | Electrodes | Workload | Interval | | |
|---|---|---|---|---|---|
| | | | Baseline | Task | Recovery |
| Theta (4 – 7) | F3, F4, F7, F8, Fz | Low | 1.35 (0.37) | 1.52 (0.33) | 1.53 (0.37) |
| | | High | 1.24 (0.37) | 1.36 (0.33) | 1.36 (0.37) |
| Alpha (8.5 – 12.5) | F3, F4, F7, F8, Fz | Low | 1.35 (0.37) | 1.52 (0.35) | 1.53 (0.35) |
| | | High | 1.24 (0.37) | 1.36 (0.33) | 1.36 (0.37) |
| Theta (4 – 7) | P3, P4, P7, P8, Pz | Low | 1.33 (0.49) | 1.53 (0.44) | 1.61 (0.49) |
| | | High | 1.23 (0.49) | 1.37 (0.44) | 1.46 (0.49) |
| Alpha (8.5 – 12.5) | P3, P4, P7, P8, Pz | Low | 1.18 (0.42) | 1.25 (0.42) | 1.48 (0.42) |
| | | High | 1.07 (0.42) | 1.06 (0.42) | 1.34 (0.42) |

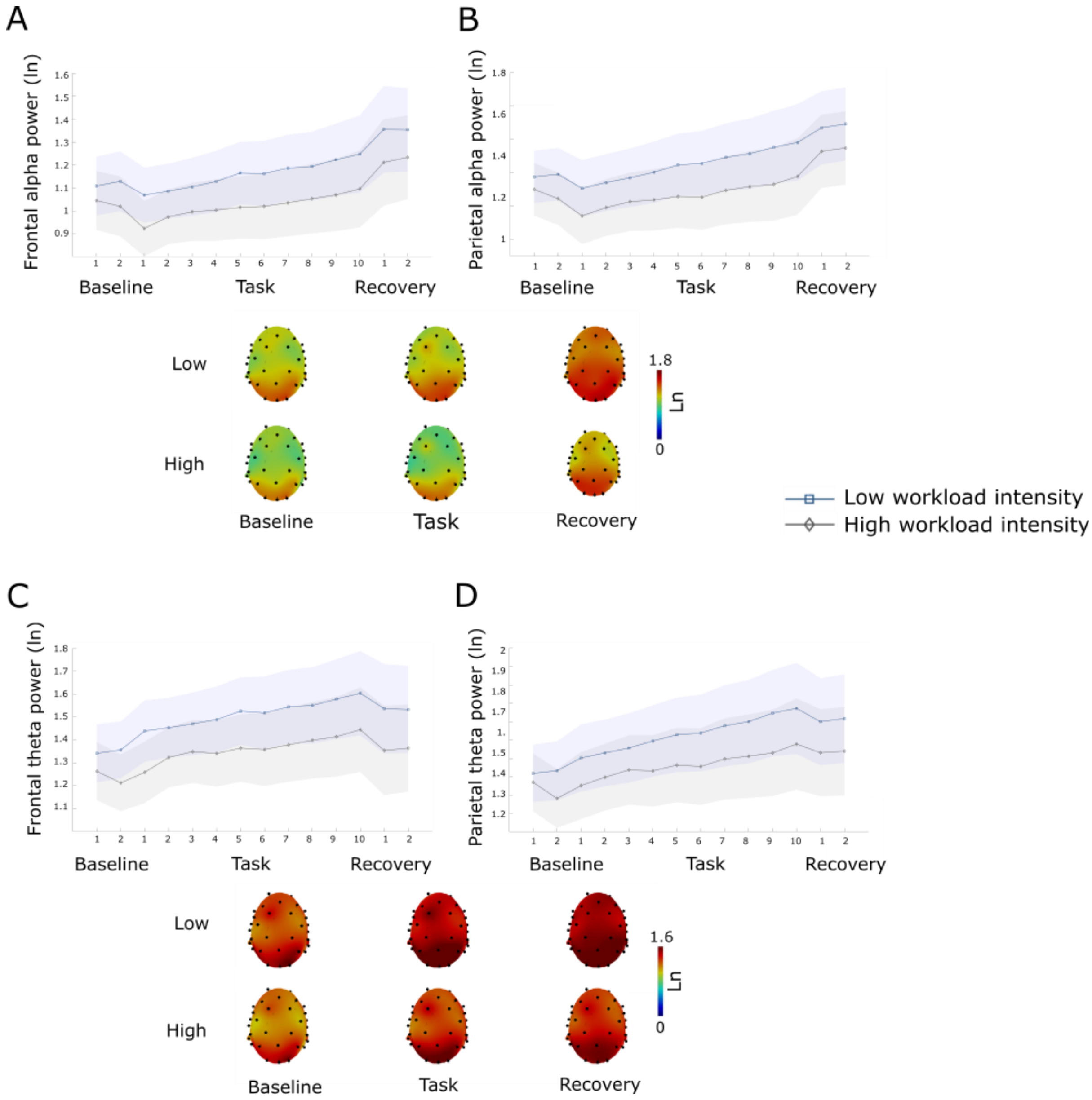



*Figure 2.7.* Grand-averaged natural logarithm transformed power for theta (4 – 7 Hz) and alpha (8.5 – 12.5 Hz) over baseline, task, and recovery periods at one-minute intervals. Shaded error bars represent the ± standard error of the mean difference. Low workload represented by the blue lines; high workload represented by the dark grey lines. **(A)** Frontal alpha power. **(B)** Parietal alpha power. Topographical plots representing alpha power averaged over baseline, during, and recovery task periods. **(C)** Frontal theta power. **(D)** Parietal theta power. Topographical plots representing theta power averaged over baseline, during, and recovery task periods.

2.4.13 Relationship between subjective ratings and cognitive performance

The linear mixed models above revealed that only subjective ratings were able to distinguish between workload during the active task period. Therefore, a participant-by-participant correlation analysis was performed on participants' cumulated subjective rating and the cumulated total score for both low workload and high workload. To ensure that the differences between low and high workload sustained for cumulated verbal rating, a paired samples *t*-test was run. The *t*-test confirmed that cumulated rating was greater during high workload ($M$ = 54.90, $SD$ = 18.07) than low workload ($M$ = 40.39, $SD$ = 16.15), $t_{(40)} = -9.31$, $p < .001$, $d_z = 1.45$. Shapiro-Wilk test of normality and visualisation of QQ plots of the unstandardized residuals revealed that normality was not violated for cumulated verbal ratings, however there was a clear outlier present in the data (cumulated score of zero for low workload and ten for high workload). As such, a Spearman's rank-order correlation was carried out. Two repetitions of the correlation analysis gave a Bonferroni corrected statistical threshold of $p < 0.025$.

The Spearman's rank correlation revealed a significant negative association between cognitive performance and verbal rating during low workload multitasking, indicating the higher the stress rating the worse the performance, but this did not survive Bonferroni corrections, $r_s = -.30$, $p = .05$. The association during high workload multitasking was not significant, $r_s = .15$, $p = .36$. See Figure 2.8 for scatter plots visualising the relationships.

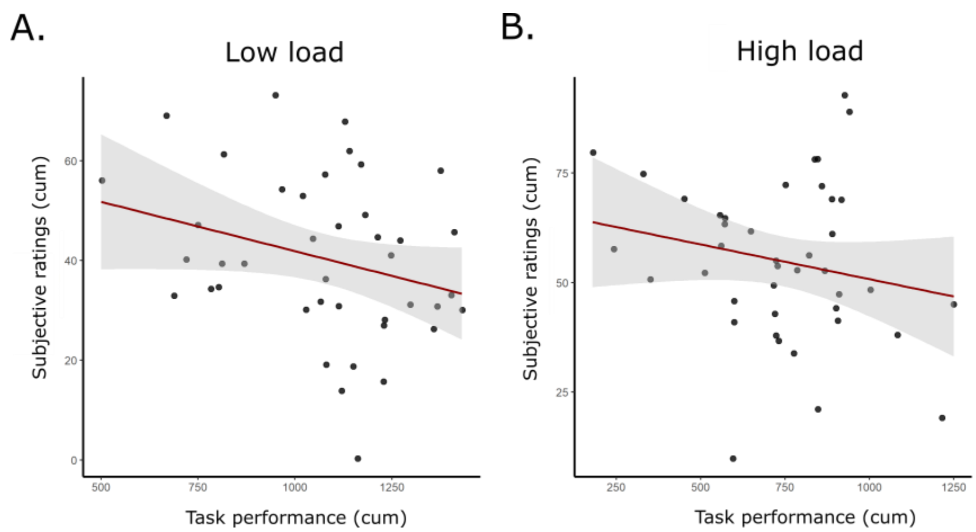



*Figure 2.8.* Scatter plots representing the relationship between cumulated subjective rating and task performance. Linear trendline depicted as a red line with shaded area representing 95% confidence interval. **(A)** Subjective ratings and task performance during low workload. This association did not survive

Bonferroni corrections (*p* = .05). **(B)** Subjective ratings and task performance during high workload. This association was not significant.

## 2.5 Discussion

This study investigated the effect of attentional load, manipulated by workload, on subjective, psychophysiological, and neural measures during multitasking. The experiment compared the effects of two workload conditions (low workload and high workload) at one-minute intervals over a two-minute baseline, ten-minute multitasking, and two-minute recovery period. Participants performed worse during high workload multitasking. Linear mixed-effects models revealed that subjective stress ratings were greater during high workload multitasking when compared to a baseline period, a recovery period, and low workload multitasking. However, subjective ratings were not associated with task performance. Heart rate was greater at the end of the task (five to ten minutes) when compared to start of the task (one to two minutes). Heart rate variability metrics (RMSSD, SDSD, pNN50) were greater during high workload when compared to low, regardless of task period (i.e. baseline, task, recovery). There were more skin conductance responses during low workload baseline, when compared to the task period. Skin conductance response amplitude was greatest at minute one, when compared to all other minutes, apart from the seventh. Skin conductance level was greatest during the first and third minute, when compared to the end of the task (seven to ten minutes). In addition, skin conductance was greater during the second minute, when compared to the rest of the task (four to ten minutes). Trapezius muscle activity was greater during the task, when compared to baseline and recovery. Muscle activity was greater during minutes three to ten, when compared to minutes one to two. EEG measures demonstrated similar effects for frontal and parietal activity. Alpha and theta activity were highest during low workload, regardless of interval. During the task, theta activity increased, while alpha activity did not differ from baseline levels. Alpha activity increased during recovery, while theta activity stayed at a similar level. These findings will be discussed below.

### 2.5.1 Subjective ratings were higher and task performance was lower during high workload

Subjective stress ratings were higher and behavioural performance was lower during high workload multitasking. Therefore, behavioural and subjective measures demonstrated an increase in cognitive demands imposed by the task across workload conditions. This is in accordance with previous research that has demonstrated performance and subjective ratings match workload conditions. Wetherell & Carter (2014) found that performance declined and perceived workload increased linearly with multitasking workload. Puma et al. (2018) also demonstrated a negative linear relationship between task performance and workload during a computer-based multitasking task. In other manipulations of workload, McKendrick & Harwood (2019) found that perceived mental demand increased when working memory load increased. In more ecologically valid tasks, the relationship between perceived workload and actual workload still stands; Liang & Pitts (2019) found that during simulated driving, perceived workload increased, and performance was worse, under increased task load. These studies support the validity of the current results.

In the above-mentioned studies, subjective measures of workload were collected via the NASA-Task Load Index (NASA-TLX; Hart & Staveland, 1988). Other techniques such as the Subjective Workload Assessment Technique are also commonly used (SWAT; Reid & Nygren, 1988). In the present Chapter, subjective ratings were obtained by requesting a number between zero and ten that corresponded to participants' "stress" level. The NASA-TLX and SWAT require participants to respond to several questions and is therefore administered once the task has ended and is completed retrospectively. In the present study, subjective ratings were collected every minute throughout the task, and therefore administering the NASA-TLX or SWAT would have been impractical and disruptive, subsequently impacting task performance. As workload is a complex multidimensional construct, only one element could be selected for participants to immediately report on. Psychological stress load was chosen as the Multitasking Framework has been shown to elicit acute psychobiological stress reactivity in controlled settings (Wetherell et al., 2017; Wetherell & Carter, 2014). However, it is important to understand the limitations of selecting one element. For example, workload elements such as psychological stress and mental effort are impacted by factors such as motivation (Rubino et al., 2009), and environment (Downey & Van Willigen, 2005), and high task demands are not directly related to high psychological stress load. For example, Mandrick et al. (2016) found that during a working memory task, unpredictable aversive loud sounds did not degrade task performance. They argued that task performance was unaffected due to malleable cognitive strategies, but at a psychophysiological cost, as indicated by measures such as increased heart rate. Therefore, participants' workload can only be inferred by current subjective ratings. In the case of the present study, it can only be inferred that subjective ratings represent psychological stress load. The correlation analysis revealed that subjective ratings were associated with task performance measures during the low workload condition. Although the effect size was moderate ($r_s$ = -.3), the relationship did not survive Bonferroni corrections. In this case, participants were not able to accurately reflect on their own cognitive performance, and so other workload elements, rather than psychological stress, may have been associated with task performance.

### 2.5.2 Electromyographic activity was highest during high workload multitasking

This was the first study to utilise the Multitasking Framework to demonstrate the impact of workload on upper trapezius muscle activity. Muscle activity increased throughout the task, and was highest at minutes three to ten, when compared to minutes one and two during baseline, task, and recovery intervals. Although electromyographic (EMG) activity was greater during high workload, a Workload by Interval interaction was not found, suggesting that muscle activity was greatest during high workload during baseline, task, and recovery. Visualisation of the results indicate that muscle activity was greater during high workload multitasking, however this did not reach significance. Therefore, trapezius muscle activity was only partly sensitive to workload.

An increase in EMG activity during high workload is supported by research studying occupational stress. Prolonged work under high workload may lead to musculoskeletal conditions particularly in the back and the shoulders as workload induces physiological changes such as increases in muscular tension. For example, muscle activity has been found to increase during sustained attention tasks (e.g. Wixted et al., 2018; Wixted & O' Sullivan, 2018), time pressure tasks (e.g. Taib et al., 2016), and psychosocial stress (e.g. Marker et al., 2017).

The present study demonstrates that workload increases trapezius muscle tension and is higher during high workload multitasking. However, the present study was unable to fully disentangle workload from baseline task measures. Yet, other research has demonstrated the sensitivity of muscle activity of the trapezius to workload conditions. For example, Deeney & O'Sullivan (2017) found that trapezius muscle activity was greater during high information processing, compared to medium and low information processing. However, participants completed the computer task while abducting their arm at a 90 degrees angle. The present study, however, measured muscle activity in general, rather than muscle fatigue. Similar to this study, Wijsman et al. (2013) administered several tasks under time pressure, but consecutively rather than simultaneously. EMG amplitude was higher during stress, when compared to rest. However, EMG metrics did not differ between stress loads. Although these results are similar to the present study, the author's tasks differed by inducing social stress via video recording. In the present study, social evaluation was kept constant throughout, as the cumulative score in the centre of the screen was consistent between low and high workload conditions. In combination, the results provide evidence that social evaluative threat and task difficulty has similar effects on trapezius muscle activity.

### 2.5.3 Heart rate and heart rate variability did not differ with increasing cognitive demands

Cardiovascular data indicated that heart rate was greater during the first minute, when compared to the second minute. However, this was regardless of whether the first minute was during the baseline, task, or recovery period. These results suggest that participants were becoming familiar and comfortable with the current environment (i.e. whether it was a resting baseline condition or the task condition). A main effect of Time also revealed greater heart rate at minutes five, six, seven, eight, nine, and ten, when compared to the second minute, suggesting that heart rate did increase, but only halfway into the task. As there was no significant difference between the baseline, task, and recovery, the Multitasking Framework did not elicit a significant increase in heart rate. Despite heart rate variability (HRV) metrics often reporting differences when heart rate does not (e.g. Fuentes-García et al., 2019), HRV was also not modulated by multitasking as although HRV was greater during low workload, this difference was regardless of interval (i.e. baseline, task, recovery).

These results are contradictory to other studies that have found an effect on cardiovascular activity during the Multitasking Framework (e.g. Wetherell et al., 2017; Wetherell & Carter, 2014). The disparities with these previous studies are two-fold. In previous studies utilising the Multitasking Framework, peak measures of heart rate were analysed. In the present study, an average of cardiovascular measures over one-minute periods were collected to understand how physiology adapted over the period of a continuous cognitively demanding task, and how this adaption impacted task performance. Therefore, it was not appropriate to measure peak reactivity. Secondly, the present study administered the Multitasking Framework for two ten-minute periods. However, previous research elicited the Framework for fifteen and twenty minutes. Therefore, task administration may have been too short. This is supported by the current study's findings that heart rate was significantly greater during the fifth to tenth minute of the task, when compared to the second minute of the task. Longer administration of the task may have demonstrated a further increase in heart rate activity.

### 2.5.4 Skin conductance responses did not differ, but skin conductance level was greater during the first minute of multitasking

Non-specific skin conductance responses (NS-SCRs) were greater during low workload baseline task period, compared to the task and recovery intervals. This is surprising, as a skin conductance response reflects a fluctuation in skin conductance level (SCL) elicited by external stimuli. As the Multitasking

Framework is a dynamic complex environment, it was expected that physiological reactivity would occur at different time points during the task. For example, when a participant loses 10 points, or when the auditory tone occurs. Therefore, it was hypothesised that skin conductance responses would occur more frequently throughout the task period. This is supported by previous research that found increases in peak physiological responses during multitasking (Wetherell & Carter, 2014). Although NS-SCR amplitude was greater during the first minute, compared to the following minutes, this was not specific to the task period. Therefore, this suggests that the task did not reliably elicit skin conductance responses.

Yet, the present study revealed tonic changes in SCL during the first minute of the task. This may therefore suggest that at immediate exposure, tonic changes in skin conductance responded to meet the demands of the environment. Rapidly, the physiological response diminished once the demands were met (Gunnar & Quevedo, 2007). SCL represents a continuous slow changing response, whereas the skin conductance response reflects a fluctuation in SCL elicited by external stimuli. Therefore, rather than the Multitasking Framework eliciting rapid phasic changes to the autonomic system, the results suggest that it elicited a tonic impact on the autonomic nervous system. However, these changes were relatively short-lasting. Performance measures suggest that the task did not become less challenging, as participants performed worse during high workload. However, the complexity of the task may have become more familiar over time, which diminished the SCL effect (Dawson et al., 2007). In line with the other physiological measures, there was no difference between workloads.

### 2.5.5 Frontal and parietal theta activity increased during multitasking, whereas alpha increased following multitasking

Findings revealed identical results between the frontal and parietal cortex in alpha and theta activity in response to multitasking. Theta activity increased during multitasking and recovery, when compared to the baseline period. In addition, theta was greater at the end of the task (ninth and tenth minute) when compared to the first and second minute, though this was regardless of interval. Theta was greater during high workload, although this was, again, regardless of interval. Alpha power was lowest at baseline versus recovery. Similar to theta, alpha power was greater during low workload.

Theta increased during multitasking, and this persisted following task cessation. This in accordance with the literature that associates theta activity with engagement and integration of task-relevant areas (Sauseng et al., 2010). Frontal theta increases are commonly associated with working memory (e.g. Sammer et al., 2007), and the task required the engagement of working memory components for successful execution. The results also reveal that theta power stayed high following task completion. It is not common to measure theta power after task completion, and so these findings provide insight into the neural after-effects of multitasking. As theta was greater during low workload, regardless of interval, theta power was not sensitive to workload. This is similar to dual-tasking research which has demonstrated increases in theta power, but no difference between task difficulty conditions (Käthner et al., 2014). However, the present study differs from Puma et al's. (2018) study that administered a similar computerised multitasking platform and found frontal theta power increased with the number of tasks, until tasks three and four, where it plateaued, noting that multitasking with four tasks was used throughout the present study. Yet, Puma et al. (2018) paradigm increased workload by increasing number of concurrent tasks, whereas this study manipulated workload by increasing task difficulty. Puma et al. (2018) also found that these differences were only apparent in medium and high performers. In sum, the most difficult variants of the task (three and four), and low performers, did not demonstrate sensitivity towards workload via theta activity. This could potentially be due to exhaustion of cognitive resource allocation. Potentially, participants may have found the low workload task difficult, and so cognitive resource allocation was already reached. However, the task performance data demonstrates a further decrement in performance during high workload, and so it can be concluded that task performance was not reflected in the theta data.

When considering alpha activity, the results show a dissimilar pattern to the literature. Alpha activity has been associated with attentional suppression across multiple sensory systems, particularly in areas associated with processing distracting information (Foxe & Snyder, 2011). A negative relationship between alpha and the blood-oxygen-level-dependent (BOLD) signal has also been demonstrated: an increase in BOLD is coupled with a decrease in alpha power (Murta et al., 2015). Therefore, during task engagement, alpha power often reduces (e.g. Foxe et al., 2014; Proskovec et al., 2019). On the other hand, various studies have demonstrated an association between an increase in alpha power and task engagement. The functional inhibition hypothesis of alpha suggests that increases in alpha are associated with the inhibition of irrelevant information (Jensen & Mazaheri, 2010). Therefore, during multitasking, an increase in alpha power might represent the inhibition of additional task processing (Puma et al., 2018). Yet, despite the clear implications of alpha power during task processing, alpha was not modulated during multitasking when compared to the baseline period. Visualisation of the data reveals a reduction in alpha power at task commencement, which steadily rises throughout the task. This may reflect the sudden recruitment of attentional mechanisms which is then confounded by a decrease in general arousal processes (Oken et al., 2006).

As this study utilised a continuous paradigm, general arousal and vigilance mechanisms may have significantly impacted neural oscillatory power, represented as an increase in alpha power.

### 2.5.6 Limitations

This study has several limitations that should be considered when interpreting its findings. It is important to understand that the Multitasking Framework task performance score is a cumulative score of all four tasks and so it is unclear what score was associated with each task separately. Therefore, participants could ignore specific tasks and only focus on the tasks they wanted to, potentially reducing the impact of high arousal multitasking. However, the Multitasking Framework does deduct points if tasks are not completed, and so ignoring specific tasks would incur a penalty and be reflected in the participant's score.

Fundamentally, an appropriate and accurate baseline is crucial. Resting physiology was measured while participants sat quietly with their eyes open. Baseline and recovery measurements were compared to task measurements, where participants were engaged with a computerised task, utilising a range of cognitive functions such as memory abilities, motor function, perceptual, and attentional mechanisms. The differences between task execution and baseline/recovery periods may have influenced the results. For example, upper trapezius activity increased during task execution. During baseline and recovery periods, participants sat still with limited movement. During multitasking, muscle activity was activated when operating a computer mouse. Therefore, differences between baseline and task may be due to task execution rather than workload. In addition, quiet, task-less baseline periods may have encouraged disruptive thoughts and mind-wandering. As such, the findings between baseline, low workload task, and high workload task, provide a greater insight into the impact of multitasking. However, it can be argued that if participants performed any task during the baseline period, workload would increase and invalidate the baseline. Therefore, in the present study, participants were asked to sit down in the same body position as the task, to make it as similar to the task period as possible. In addition, as stated by Laborde et al. (2017), following standardised baseline measurements will make it possible to compare results across research. Therefore, the present study followed their recommendations for baseline measurement in cardiovascular research (Laborde et al., 2017). Regarding muscle activity, all participants utilised their right hand for mouse control. EMG activity of the left trapezius muscle only was analysed, to ensure that movements of the right arm had limited impact on measurements. This approach is similar to previous research (e.g. Wijsman et al., 2010).

Physiological responses may be affected by the anticipation of an upcoming task as excessive levels of apprehension have been found during anxious anticipation (Grupe & Nitschke, 2013). Yet, it is important to take baseline measures at a similar time as the task, as otherwise many factors may influence the baseline measure, such as time of day, food, or sleep (e.g. van Eekelen et al., 2004). Utilising a recovery period provides insight into baseline measures which are not confounded by anticipation effects. The impact of this can be seen with the skin conductance response frequency data. There were a greater number of skin conductance responses at baseline, when compared to the task, and recovery period. An anticipation effect may have instigated this increase (Dawson et al., 2007).

### 2.5.7 Conclusion

Subjective ratings successfully disentangled low and high workload multitasking. Cardiovascular, electrodermal activity, muscle activity, and neural activity were not sensitive to workload changes. Cardiovascular data provided inconclusive results which emphasise the need for baseline and recovery measures, to ensure differences between workload conditions are relative to baseline. Tonic increases in skin conductance rapidly diminished after the first minute of the task, suggesting task demands were met promptly. Non-specific skin conductance responses demonstrated an increase at baseline, indicating an anticipatory effect. Frontal and parietal theta activity engaged task-relevant areas which persisted following task cessation. Frontal and parietal alpha task-related activity did not differ to baseline, yet alpha power increased following the task which might reflect a global decrease in arousal and vigilance mechanisms.

The findings of this Chapter informed future studies described in subsequent Chapters in two ways. Firstly, in this experimental design, the effects of workload might be too small to reliably impact autonomic and central nervous system mechanisms. Yet, the study did reveal a lasting effect of multitasking on neural mechanisms. Therefore, this finding inspired future work described in Chapter 3 which employed another manipulation of multitasking, namely an adaptive version of the task-switching paradigm, to attempt to unravel the negative lasting impacts of multitasking. In addition, Chapter 3 employs a different methodological design (event-related) to increase the signal-to-noise ratio, and remove within-subject variation, to separate workload conditions. Secondly, this Chapter also revealed insight into suitable measures of sustained attention during applied contexts, informing future studies described in Chapters 5 and 6. As self-report ratings were sensitive to workload changes, Chapters 5 and 6 also included self-reports where possible. In addition, as tonic changes in skin conductance were sensitive to the immediate onset of the task, skin conductance level was subsequently measured to detect fluctuations in alerting and orienting during naturalistic studies.

# 3.0 The Effect of Prior Attentional Load on Task Performance and Neural Oscillations During a Continuous Tracking Task

## 3.1 Overview

Based on the findings in Chapter Two, event-related alpha and theta oscillatory activity were measured during a continuous pursuit tracking task that was preceded by a visual search task of either low or high perceptual load. This scenario is crucial in applied contexts, when sudden and critical alerts require a user to switch tasks accurately and efficiently. For example, physicians responding to a major incident alert, or drivers during takeovers in semi-autonomous driving. These sudden but vital task switches have life-threatening consequences and are made more challenging when humans have the potential to engage in non-related tasks, such as reading during periods of automated driving. Does the complexity and demand of the previous task (i.e. how engaging the book is/attentional load) have an impact on secondary (i.e. driving/tracking) task performance?

## 3.2 Introduction

Task-switching is an important aspect of multitasking. Rapidly reorienting attention to relevant sensory stimuli involves the recruitment of attentional and goal-related systems engaging task-related networks while disengaging task-irrelevant networks (Petersen & Posner, 2012). Adopting a task-set selects the appropriate task-relevant processes which facilitates successful attention reorientation (Rogers & Monsell, 1995). For example, when reading a book, the movement of eyes from left to right and the act of reading adopts a task-set that successfully selects processes such as ocular motor control, working memory, and semantic processing. The process of changing the existing task-set to perform a different task is termed task-set reconfiguration (Rogers & Monsell, 1995). Behavioural studies have consistently demonstrated that task-set reconfiguration results in increased reaction times and/or reduced accuracy (for reviews, see Jamadar et al., 2015; Monsell, 2003). This *switch cost* can be reduced if the upcoming task is predictable or if a cue indicates the upcoming task (Meiran, 1996; Rogers & Monsell, 1995). If the cue-stimulus interval is long, e.g. 1000ms, switch costs are reduced further (Meiran, 1996; Rogers & Monsell, 1995). Not only are switch costs impacted by temporal factors, the physical properties including the congruency of target items (e.g. Schneider, 2018), the affective salience of the stimuli (e.g. Reeck & Egner, 2014), and the level of arousal (e.g. Solano Galvis et al., 2010) also modulate switch costs. Yet, switch costs are rarely eliminated. These results suggest that exogenous processes related to the task stimulus are important for successful task switching and may impact a carryover effect from the previous task-set, termed *task-set inertia* (Allport et al., 1994).

### 3.2.1 Neurophysiological correlates of task switching

Task switching is considered a distinct executive process that relies upon a top-down control system. This system requires activation of working memory in accord with the shifting task-set, activation of the upcoming task-set, while supressing previous task-set activity (Mayr et al., 2014). Behavioural measures such as average reaction time and accuracy provide relatively crude information and are unable to offer direct insight into the temporal evolution of the top-down processes involved. The excellent temporal resolution of non-invasive electrophysiological data and spatial resolution of functional magnetic resonance imaging (fMRI) techniques can therefore be used to study the task-switching process at the neural level and uncover the networks involved in the preparation and execution of the behaviour.

By comparing brain regions that are activated during a switch trial (i.e. two different tasks) and a repeat trial (i.e. the same task), studies have found that the prefrontal cortex (e.g. Dove et al., 2000), the anterior cingulate (e.g. Liston et al., 2006), and the posterior parietal cortex (e.g. Sohn et al., 2000) are involved during task switching. These results implicating a distributed network of fronto-parietal regions are supported by a meta-analysis which revealed that general task-switching processes consistently activate areas in the inferior frontal junction and posterior parietal cortex (Kim et al., 2012). This is in accord with the task-set reconfiguration model, which emphasises the importance of top-down executive control functions. By investigating domain-specific brain regions of interest, researchers have provided evidence for the interference of a previous task-set on switch costs. Yeung et al. (2006) found fronto-parietal areas associated with the former task were still activated during the secondary task. In addition, the blood-oxygen-level-dependent (BOLD) signal from task-irrelevant brain areas positively correlated with behavioural switch costs, such as increased reaction times and error rates. In line with task-set inertia accounts, these results revealed that interference from the previous task-set was associated with poorer task performance.

Research focusing on oscillatory activity has demonstrated mechanisms related to suppression of the previous task-set. Alpha activity has been associated with attentional suppression across multiple sensory systems, particularly in areas associated with processing distracting information (e.g. Foxe & Snyder, 2011). Recent research has demonstrated stronger and earlier alpha event-related desynchronisation (i.e. reduced alpha power; ERD) across fronto-central and parieto-occipital cortices during a task switch, when compared to repeat trials (e.g. Foxe et al., 2014; Proskovec et al., 2019). Proskovec et al. (2019) demonstrated that greater alpha desynchronisation positively correlated with smaller behavioural switch costs. As alpha suppression has been shown to indicate task engagement (see Chapter 1.2.4), this evidence suggests that decreases in alpha activity might lead to improved readiness for a task switch.

As an executive process, mechanisms of task-set reconfiguration include the activation of task goals to direct behaviour. Therefore, the relationship between goal directed oscillatory mechanisms and switch costs have been investigated. Cooper et al. (2017) presented a cued task-switching paradigm and found increased fronto-parietal theta activity during the cue period. In addition, mid-frontal theta has been found after target onset (Cooper et al., 2017; Cunillera et al., 2012). Cunillera et al. (2012) found a positive correlation between P3b amplitude and frontal midline theta power for switch cues. As the P3b is related to stimulus context updating processes (Polich, 2007), together, these findings suggest that frontal-parietal theta activity plays a significant role during goal updating processes. Mid-frontal theta is also regularly found to correlate with working memory load and performance (e.g. Brzezicka et al., 2018) which also supports theta having a significant role during successful task-switching.

### 3.2.2 Relating neurophysiological processes during task switching in applied contexts

While the aforementioned studies have provided insight into the structural and functional networks involved in task-switch performance, they are limited in several ways. Task switching is often cited as a significant challenge in everyday life as scenarios requiring individuals to switch between different activities can have a critical impact on safety, health, and wellbeing. Yet task-switching paradigms employ short lasting trials (i.e. two seconds) and simple stimuli that share the same features across conditions (i.e. letters). Although it is important to control basic features to unpick the processes involved, it makes it difficult to transfer the findings to applied contexts where multiple processes interact and communicate with one another (also discussed in Chapter 1.3.3). In addition, the impact of task-set inertia is transient in typical task-switching paradigms, yet sustained effects have been found depending on the task type (i.e. if task stimuli are similar between both tasks; Waszak et al., 2003). Therefore, it is not clear how long task-set inertia effects may last in a real-world scenario that is not limited by short trials and simple visual stimuli. Investigating task-switching processes over a longer time period during a sustained task requiring engagement of multiple processes including motor action is needed.

An extensive network of regions have been implicated during a task switch, yet these areas overlap with many cognitive processes and subprocesses. For example, many of the brain regions implicated during a switch trial are part of the attention reorientation systems: the dorsal attention network (DAN), and the ventral attention network (VAN; Petersen & Posner, 2012). Oscillatory dynamics are also similar between attention-related and switch-cost related networks, including the inhibition of parietal alpha activity and the increases in frontal theta activity (see Proskovec et al., 2018, 2019). Therefore, the areas and mechanisms that are typically defined as ‘task-relevant’ or ‘attentional’ are likely to include areas that are important for a multitude of higher-order processes such as vigilance and response planning (Fortenbaugh et al., 2017). It is therefore important to appreciate that task switching in applied contexts must be informed by many areas of literature including attention-reorienting and working memory.

Although the brain regions involved in task-switching behaviour have been well characterised via fMRI investigations, less is understood about the dynamics of oscillatory activity. The majority of previous studies have focused on oscillatory dynamics during the preparatory period, rather than the secondary task period (e.g. Cooper et al., 2016, 2017). Yet, it is important to understand the impact of a previous task on consequent task performance. When the secondary task period has been investigated, analogous simple stimuli have been employed (e.g. Barceló & Cooper, 2018; McKewen et al., 2020). Therefore, to understand the effects of a previous task on new task performance, research is needed investigating performance and neural oscillatory dynamics during the secondary ‘main’ task.

### 3.2.3 Experiment rationale

The goal of the present study was to utilise EEG to investigate the oscillatory dynamics underlying a visuomotor task when preceded by a visual search task differing in perceptual load (Lavie, 2005, 2010; for a review, see Murphy et al., 2016). This was designed to understand how long activated task-sets persist over time and interfere with the new task-set configuration, beyond simple reaction time and accuracy effects in transient tasks, as previously described in the literature (for reviews, see Jamadar et al., 2015; Monsell, 2003). This is particularly important in applied contexts, such as the task of taking back control of the vehicle following a period of non-driving during semi-autonomous driving. As the driver will have the potential to engage in non-driving related tasks such as reading, this is a critical task with life-threatening consequences. Does the context and demand of the previous task (i.e. how engaging the book is) have an impact on secondary (i.e. driving) task performance?

A visual search task was chosen for several reasons. Firstly, visual search is not confined to the laboratory. For example, we spend time looking for a friend in a crowded bar, looking for our keys at home, or looking for milk in the refrigerator. These examples of visual search processes require the interaction of cognitive processes including perception, attention, and working memory. Secondly, the literature demonstrates that perceptual load has a performance effect over short timescales (i.e. one second), yet it is not clear how long the short-term effects of perceptual load last. Finally, visual search tasks have been used as important experimental manipulations of attentional demand. Therefore, the neural and performance effects of visual search are well documented (e.g. Chelazzi et al., 1993; Eimer, 2014; Wolfe & Horowitz, 2004). For example, searching for a target item amongst distractor items result in performance costs when the number of distractors increase (for a review, see Murphy et al., 2016). Therefore, a task of higher perceptual load, i.e. a higher number of distractors, requires greater task engagement reducing spare attentional capacity. This indicates that it may be more demanding to switch to a distinct secondary task. If performance of the secondary task is not affected, then task-switch neural processes may compensate and increase during the secondary task. If performance deficits are found, then neural processes may not be strong enough to engage the required attentional resources needed for optimal performance. This also provides the motivation for measuring neural oscillatory activity; attentional and general cognitive demand associated with task switching indexed by neural oscillations such as alpha and theta may be greater following high workload if behavioural effects are not found.

The secondary task was a continuous pursuit tracking task. This task is a visuomotor task that requires the participant to track a moving target with a cursor. A pursuit tracking task can be related to everyday sustained visuomotor tasks such as driving. Broeker et al. (2020) recently found that detriments to performance in both pursuit tracking and simulated driving were comparable, suggesting similar cognitive costs to networks involved in attention, motor control, and eye–hand coordination. However, it is important to recognise that driving includes both controlled and automatic processes, whereas tracking in an experimental condition will only include controlled processes. Experienced drivers on a familiar route will often find driving easy without any effort, yet the controlled nature of the tracking task requires effortful control throughout. In addition, oscillatory mechanisms have been shown to be similar across sustained attentional tasks, including a tracking task and simulated driving (Huang et al., 2007). Thus, the tracking task is advantageous over other laboratory-based measures as it allows manipulation of multiple pathways and can therefore be associated to everyday tasks such as driving. Secondly, although the tracking target moves in a random direction, it continues on the same trajectory. This is similar to activities such as driving (i.e. following the lane ahead; Ramachandran & Anstis, 1983). Therefore, an important aspect of visual attention in applied and tracking task contexts deals with the relationship between bottom-up and top-down processes, including how goal directed behaviour can impact visual behaviour (Betz et al., 2010). Finally, both transient and sustained changes in performance can be recorded. Therefore, pursuit tracking can capture small lapses in accuracy and reaction time throughout a continuous task. Sustained changes in performance are also important to evaluate because fluctuations in reaction time and accuracy during everyday activities, such as not responding quickly enough to a hazard during driving, can be fatal.

The hypotheses are related to neural mechanisms that vary with attentional load during a visuomotor task. Therefore, the hypotheses are unable to identify neural mechanisms involved specifically in task switching, as the activity involved will include other higher order processes such as response planning, as discussed above. Nevertheless, the task switching and attentional load literature can provide insight into expected behavioural and neural differences (e.g. Proskovec et al., 2018, 2019). In accordance with the literature (Lavie, 2005, 2010), visual search high load will reduce task performance, as demonstrated by reduced accuracy and increased reaction times. As the visual N1 event-related potential (ERP) component is related to sensory processing, the N1 should increase during high load as there will be a greater number of distractors in the visual field. During low load, participants should be more likely to select the target, eliciting a greater P3b. Given that high load demands greater attention and cognitive control processing, parieto-occipital alpha (increased alpha event-related desynchronisation (ERD)) should be lower, and frontal theta (increased theta event-related synchronisation (ERS)) should be greater, during the high load visual search task.

The pursuit tracking task over several seconds produces two measures of immediate measures of task-switch readiness: time to first mouse movement, time to move the mouse inside the target, and two continuous measures from throughout the tracking task: the number of deviations away from the target, and the total duration of deviations. As reliable performance deficits have been found during visual search high load, reaction times during the tracking task following high load should be negatively impacted, as demonstrated by an increase in time to first mouse movement and time to the target. If carryover effects are prolonged, then the number of deviations and the total duration of deviations will be greater following a visual search task of high load. If there is less spare attentional capacity and more interference from the previous task-set following the visual search task of high load, then demands on attention and cognitive control should be greater during, or at least at the start of, the tracking task. If so, this might be reflected by lower parietal alpha (increased alpha ERD) and greater frontal theta (increased theta ERS) during the tracking task. The time course of such oscillations will be inspected to investigate the timing of task-switch effects. As lateralisation effects have been found for attention-related alpha during continuous motor tasks (e.g. van der Meer & van der Weel, 2017), alpha should be lower (increased alpha ERD) in the contralateral (relative to movement) parietal hemisphere.

## 3.3 Method

### 3.3.1 Participants

Forty healthy right-handed adult participants (twenty-four females, sixteen males, mean age ± SD = 21.54 ± 4.59 years, range 18-40 years) participated in this study. Data from one participant were excluded due to EEG recording failures. Two participants demonstrated excessive artefacts in their EEG trace and were subsequently removed. One participant withdrew during the experiment. Data from the remaining thirty-six participants (twenty females, sixteen males, mean age ± SD = 21.43 ± 4.53 years, range 18-40 years) were included in all analyses. All participants considered themselves healthy as reported in self-reports. Due to the nature of the experiment, individuals with hypersensitive skin or skin allergies, a neurological condition, or with uncorrected vision or hearing were excluded. The majority of participants were recruited through the University of the West of England's Psychology Participant Pool. Psychology students who participated were eligible to receive three and a half course credits for completing the study, in keeping with standard practice for psychology undergraduate course participation. All participants gave written informed consent and were fully debriefed at the end of the study. Ethical approval was obtained by the Faculty of Health and Applied Sciences University of the West of England Research Ethics Committee (HAS.18.10.050).

### 3.3.2 Stimuli

Visual stimuli were created and presented using Psychtoolbox-3 (http://psychtoolbox.org/; Brainard, 1997; Pelli, 1997), running in MATLAB (MathWorks, Natick, MA). The stimuli were shown on a Dell 17-inch LCD monitor with a 60 Hz refresh rate, display resolution 1280x1024, and 57 cm viewing distance. See Figure 3.1 for a graphical illustration of the experimental stimuli. A white fixation dot was presented at the centre of a black screen for 1500 ms (+/- 250 ms) at the start of each trial. At the start of the visual search task, a white fixation dot was presented for 500 ms. A circular array of six letters were presented for 150 ms. Each letter was equally spaced in a circle of 120° radius that was centred at fixation. All letters were of the same size and were coloured off-white (RGB values: 204, 204, 204). Participants were instructed to discriminate between the target letters 'M' and 'N' as quickly and accurately as possible. Each target letter to appear was random, but equally likely to appear at one of the six locations. Responses were made using their dominant hand for both letters (index finger for N, middle finger for M). All participants chose to use their right hand, regardless of handedness. Participants were instructed to maintain central fixation during the stimulus presentation to minimise eye movements and preserve visual stimulation consistency. The 150 ms duration of the presented stimuli was sufficiently brief so as not to allow for a saccade to the target. Participants had 1.5 s to respond by selecting the appropriate key. Once the participant had responded, the fixation dot came off the screen for 100 ms to provide feedback to the participant. This process (fixation dot: 500 ms, visual search: 150 ms, response: < 1.5 s) was repeated randomly between four to seven times.

Once all visual search stimuli were presented, a 500 ms fixation dot was displayed, to ensure the participant did not anticipate the task switch. Next, the screen changed to white and an auditory beep was presented for 100 ms simultaneously (600 Hz, volume 0.5 normalised). This acted as a visual and auditory cue, indicating the start of the secondary task. Auditory stimuli were generated using the Psychtoolbox function *Beeper*. The secondary task was a pursuit tracking task that lasted for 5 s. On a black background, a red target circle centred in the middle of the screen moved in a random direction but followed a straight trajectory. When the target reached the edge of the screen, it would bounce off at 180° before following a straight trajectory. The cursor was displayed as a white cross and appeared at the centre of the screen. Participants were instructed to grab the mouse and keep the cursor in the centre of the target. They were told to be as quick and accurate as possible. In summary, a trial can be considered a 'composite' trial made up of an initial fixation cross, four to seven visual search presentations (of the same perceptual load), a visual and auditory cue, followed by a 5 s tracking task.

Each participant performed low and high perceptual load trials, which were mixed between trials to prevent fatigue effects between conditions. Only the visual search task (the primary task) differed in load. In the low load condition, the distractor letter was 0. In the high load condition, the distractor letters were H, K, Y, W, and Z. The tracking task (the secondary task) was consistent between the low and high load conditions. The stimuli were presented in four blocks of 48 trials. Each trial lasted approximately 10 s. Each block lasted ~12mins. In total, 192 trials were presented: 96 trials low load, and 96 trials high load.

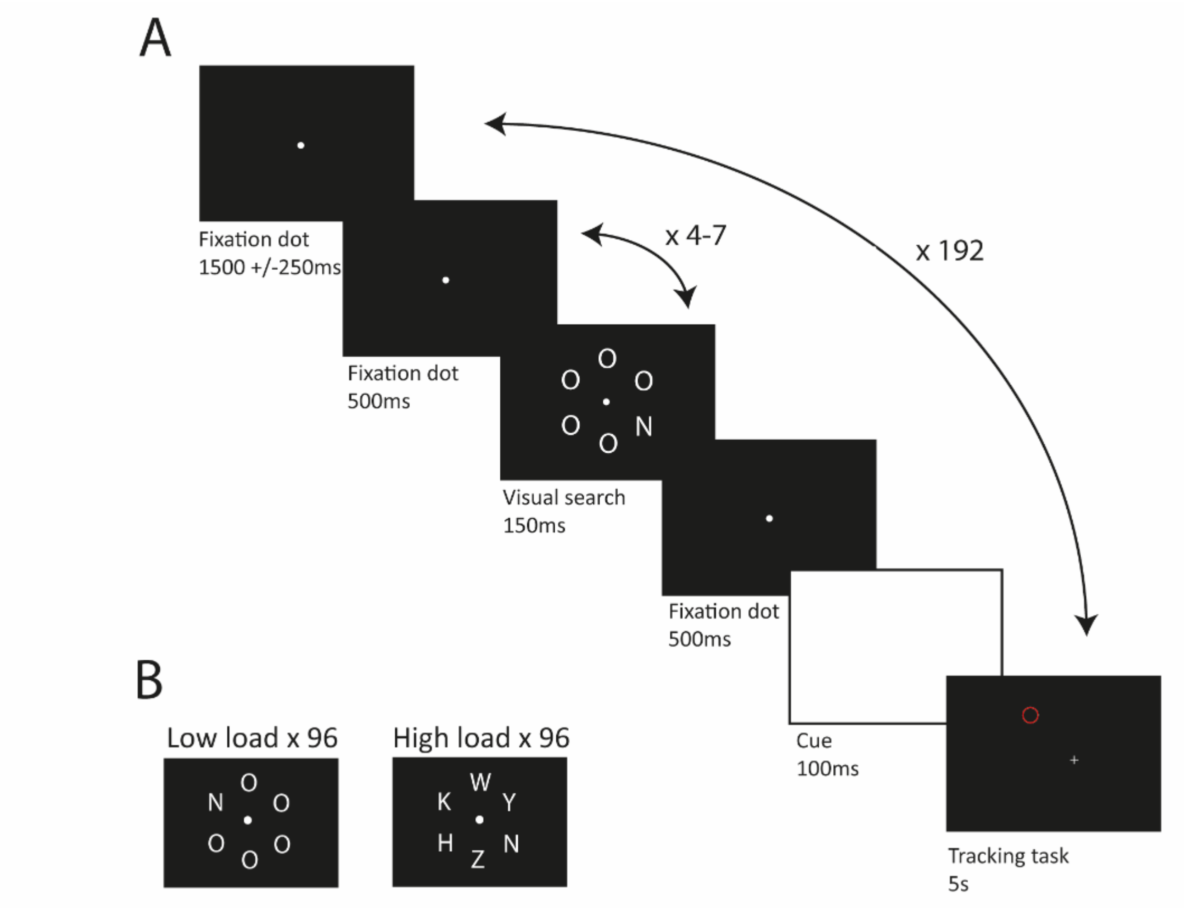



*Figure 3.1.* Schematic depiction of the experimental paradigm. **(A)** An example composite trial. An initial central white fixation dot was presented for 1500+/-250 ms. A central white fixation dot was presented for 500 ms followed by the visual search presentation for 150 ms. The fixation dot and visual presentation were repeated randomly between four and seven times. A fixation dot was presented for 500 ms, followed by an auditory and visual cue for 100 ms. A 5 s pursuit tracking task consisted of a red target circle and a white cross as the cursor. 192 trials were presented in total. **(B)** Each composite trial differed between visual search low load or high load. The visual search presentation consisted of one target letter (M or N), with no distractor items (O) or five distractor items (W, X, K, Z, Y), and was considered low and high perceptual load respectively. Each condition was repeated 96 times.

### 3.3.3 EEG recording

Continuous EEG was recorded using BrainVision actiCHamp 32-electrode EEG system with active electrodes. The ground electrode was located at Fpz (10-5 electrode system: Oostenveld & Praamstra, 2001). The online sampling rate was set at 1000 Hz and were online referenced to Cz position. Impedances of all channels were adjusted to below 25 kOhms. Continuous EOG was recorded using the same system with passive electrodes connected to the actiChamp amplifier. Two pairs of electrodes were placed at the outer canthi of both eyes, and above and below the right eye to record horizontal and vertical eye movements, respectively. An additional ground electrode was placed on the left side of the forehead.

### 3.3.4 Protocol

Participants received the information sheet via email immediately after signing up to the study. On arrival to the laboratory, participants were reminded of the content of the information sheet, introduced to the equipment, and asked whether they had any concerns or questions. Then they signed printed copies of the consent form and filled in the demographic questionnaire. They were provided with verbal instructions for the cognitive task and completed a practice run consisting of eight randomised trials. The principal investigator observed their responses and made an informed decision whether they showed an objective understanding of the task. All participants showed an understanding of the task rules after the practice run.

Once the EEG set up was complete, all electrophysiological signals were checked visually before commencement of the experiment. A within-subject protocol required participants to switch between visual search presentations and a pursuit tracking task over four experimental blocks consisting of 48 trials. In each block, participants completed 24 low load trials and 24 high load trials, resulting in 96 low load and 96 high load trials collected. Rest breaks occurred between blocks to allow participants to move freely and repose. At the start of the experiment, at the end of the experiment, and after each rest break, resting-state potentials were recorded during a two-minute eyes open and two-minute eyes closed procedure. This resulted in five two-minute eyes open and five two-minute eyes closed recordings. Overall, the task lasted ~1 hour including eyes-open and eyes-closed recordings. Figure 3.2 displays the experimental procedure timeline.

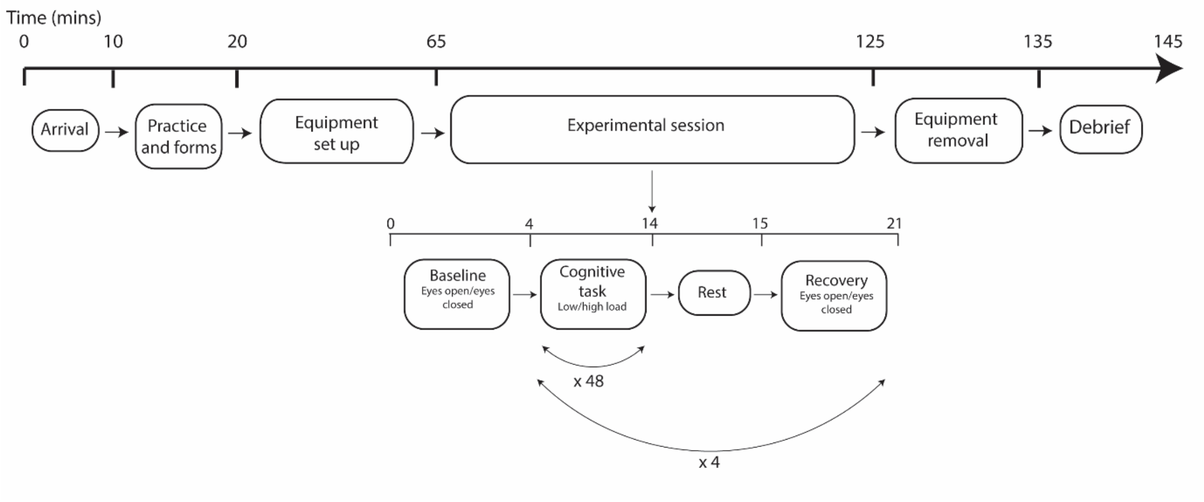



*Figure 3.2.* The experimental procedure. A within-subject protocol required participants to switch between visual search presentations and a pursuit tracking task over four experimental blocks consisting of 48 trials. In each block, participants completed 24 low load trials and 24 high load trials, resulting in 96 low load and 96 high load trials collected. Rest breaks occurred between blocks to allow participants to move freely and repose.

## 3.3.5 Pre-processing

### *3.3.5.1 Behavioural data*

Pre-processing of behavioural data was undertaken using custom written scripts in MATLAB (MathWorks, Natick, MA). Behavioural responses were collected for the visual search task and the tracking task. For the visual search task, responses were considered accurate if the participant responded with the correct letter (M or N) of the target stimulus between 200 and 1500 ms after the onset of the array. No participants responded before 200 ms. Mean accuracy was calculated (correct trials divided by total number of trials). Response time was recorded as the time between the onset of the array and the button press for each correctly reported trial.

Four measures were derived for the tracking task. Two instant measures of performance were calculated via a reaction time and accuracy reaction time measure. Reaction time was recorded as the time between the onset of the tracking target and the time the cursor moved (indicated by a change of X and Y coordinates of the cursor). Accuracy reaction time was recorded as the time between the cursor first moving and the cursor entering the target circle. Two continuous measures of performance were recorded: the frequency of deviations from the target and the total time of deviations from the target. Frequency refers to the number of deviations made where the cursor went outside the circle. Total time was recorded for each of the deviations. Means were derived for all measures.

*3.3.5.2 EEG data*

Pre-processing of EEG data was performed using EEGLAB (Delorme & Makeig, 2004) and custom-written scripts in MATLAB. Firstly, the data were downsampled to 500 Hz using the resample function. Next, the data were bandpass filtered using an FIR filter between 0.1 to 100 Hz. Data were then epoched into two separate sets: visual search and tracking task.

For the visual search task, the data were epoched from -1.5 s to 2 s in relation to visual search onset (0 s). The data were linearly baseline corrected from -200 ms to 0 ms. Only correctly responded to trials were analysed, and so trials with an incorrect or no response were removed from subsequent analyses. Only correctly responded to trials were analysed to ensure that the EEG activity was related to low perceptual load or high perceptual load and was not contaminated by an incorrect response. Manual trial rejection was undertaken to extract major sources of artefacts. Next, ICA was computed for further artefact rejection. ICA was computed on each dataset and the topographies and waveforms of the 30 leading components were visualised. One or two leading components that showed a clearly non-cortical origin (eye blinks) were removed. The newly generated data were compared to the previous data to ensure only eye movements were removed, and the signal was left. Once artefacts were removed, the data were re-referenced to the average of all EEG channels.

For the tracking task, the data were epoched from -1.5 s to 6.1 s in relation to the cue onset (0 s), resulting in the tracking task onset at 100 ms. The data were linearly baseline corrected from -200 ms to 0 ms. Manual trial rejection and ICA was undertaken as described above to extract major sources of artefacts. Once artefacts were removed, the data were re-referenced to the average of all EEG channels.

*3.3.5.2.1 Event-related potentials (ERPs)*

To capture visual search event-related potentials (ERPs), a low-pass filter of 35 Hz was applied. The data were then epoched from -200 ms to 1000 ms, and condition averaged to create subject specific ERPs. This resulted in subject specific ERPs for the visual search low load and the visual search high load. Electrodes and time windows for ERP analyses were selected based on visual inspection of the corresponding condition grand-averaged waveforms (also known as the grand-grand average waveform; Kappenman & Luck, 2016). The grand-grand average collapses the waveforms from both conditions and therefore minimises any bias towards the largest difference between groups, as this should not be visible in the grand-grand averaged waveform (Kappenman & Luck, 2016). Time windows were also advised by previous research (stated below).

*P3b.* To investigate the impact of perceptual load on attentional allocation, mean amplitude of the P3b was extracted for each visual search presentation during low and high perceptual load. Maximum amplitude of P3b has typically been reported at the mid-line channels of scalp surface, and visualisation of grand-grand averaged ERP waveforms confirmed this. Therefore, channel Pz was used to calculate the mean amplitude. To identify the appropriate time window of interest, a grand-grand average was visualised. Similar to previous research, a latency window of 300 – 450 ms was used (e.g. Barceló & Cooper, 2018).

*N1.* To investigate the impact of perceptual load on perceptual processing, mean amplitude of the N1 component for each participant was taken in a 50ms latency window around the peak of the N1 as visualised in the grand-grand averaged ERP waveform (peak = 170 ms; range = 145 – 195 ms). Data from electrodes O1, O2 and Oz were extracted and averaged together.

*3.3.5.2.2 Spectral analyses*

To capture spectral oscillatory data, the re-referenced data were epoched into two separate datasets: the visual search task, and the tracking task. The epoched data were decomposed into a time-frequency representation with linear scaling between 3 and 100 Hz from fast Fourier transform and via Morlet complex wavelet convolution, followed by the inverse fast Fourier transform. Cycles were increased with frequency, starting at 3 and going up in steps of 0.5. This ensured that frequency and time resolution was constant, and that lower and higher frequencies were decomposed. The wavelet transform was performed for each trial. Power ($\mu V^2$) of oscillatory activity was computed. Therefore, this measure of signal amplitude in single trials reflects the total amplitude for a certain frequency range, irrespective of whether it is phase-locked to the stimulus or not. It includes both evoked as well as induced activity. In order to remove scale differences between individuals, all power values in the time-frequency representation were normalised by decibels (dB) to the baseline power. For the visual search task, baseline power was computed as the average power from -400 to -200 ms pre-stimulus at each frequency band (db power = $10\log_{10}$ [power/baseline]). Unlike ERP analyses, where the baseline period ends at time 0, an optimal baseline period for time-frequency analyses ends before time 0. This is because decomposition of the signal to the frequency domain results in temporal leakage of trial-related activity. Therefore, if activity of interest occurs 50 ms after stimulus onset, during decomposition, when temporal smoothing happens, it might transpire before time 0 in the final time-frequency spectrogram (Cohen, 2014). It is therefore common to have a baseline period ending between -200 ms or -100 ms before stimulus onset (e.g. Cooper et al., 2017; Simonet et al., 2019; Yordanova et al., 2001). For the tracking task, the baseline power was computed as the average power from -500 to -100 ms pre-stimulus. A condition-specific baseline was used as differences were expected in the baseline period per condition, due to interference from the primary task (Cohen, 2014). Lastly, the absolute values of the resulting transforms were trial averaged to provide a time-frequency decomposition for each participant over every electrode.

A condition (low and high load) grand-averaged time-frequency decomposition was visualised over every electrode for both the visual search and tracking task. Similarly to the ERP analyses, the condition grand-averaged time-frequency decomposition collapses the time-frequency decomposition from both conditions and therefore minimises any bias towards the largest difference between groups, as this should not be visible in the collapsed-averaged time-frequency spectrogram. Approximate time-frequency regions were defined based on a priori hypotheses and previous research (see below). Then, visualisation of the collapsed-averaged spectrogram provided an exact time window and exact frequency range to extract data from. The activity from all pixels in a region of interest were averaged for each participant. This approach is designed for hypothesis testing and is guided by previous research and the characteristics of the results (Cohen, 2014). See Appendix 3.1 and Appendix 3.2 for condition grand-averaged time-frequency spectrograms.

*Visual search.* Regions of interest were defined for alpha (8.5 – 12.5 Hz) and theta (4 – 7 Hz) power. Frontal activity was averaged over electrodes Fz, FC1, FC2 between time 200 – 600 ms for alpha, and 200 – 600 ms for theta. Parietal-occipital activity was averaged over electrodes O1, O2, P3, P4, P7, P8 between time 200 – 600 ms for alpha, and 100 – 300 ms for theta.

*Tracking task.* Regions of interest were defined for theta (4 – 8 Hz), alpha (8.5 – 12. 5Hz) and beta (15 – 25 Hz). Frontal activity was averaged over electrodes Fz, FC1, FC2. For frontal theta, data were averaged between time 100 – 400 ms. For frontal alpha, data were averaged between 350 – 5100 ms. Parietal activity were averaged over contralateral electrodes (P3, P7) and ipsilateral electrodes (P4, P8). For alpha, data were averaged between 250 – 5100 ms, whereas theta was averaged between

150 – 450 ms. Motor-related activity was extracted from electrode C3 between 150 – 5100 ms for beta activity, and between 250 – 5100 ms for alpha power.

### 3.3.6 Statistical analyses

All statistical analyses were performed using IBM SPSS Statistics for Windows, version 25 (IBM Corp., Armonk, N.Y., USA). All *t*-tests undertaken were two-tailed. For two-way ANOVAs, assumptions of sphericity were not violated as indicated by Mauchly's test. Effect size was reported as partial eta squared ($\eta_p^2$) for two-way ANOVA significant results (Cohen, 1988). Cohen's $d_z$ was reported for paired-samples *t*-tests.

*Behavioural analyses.* Paired samples *t*-tests were used to compare accuracy and response time measures between visual search low and high load conditions. For the tracking task, paired samples *t*-tests were undertaken to compare reaction time, time to target, frequency of deviations, and total time of deviations, between the preceding low load and high load conditions.

*EEG analyses.* For ERP components, amplitude was compared between low and high load with paired samples *t*-tests. For visual search spectral analyses, normalised power (decibel) was compared between low and high load with paired sampled *t*-tests.

For tracking task spectral analyses, paired-samples *t*-tests were undertaken to understand the effects of load on evoked midline frontal theta activity, frontal alpha activity, and motor-related beta and alpha activity. A two-way repeated measures ANOVA was performed to understand the impact of low and high perceptual load on parietal alpha activity and parietal theta activity during the tracking task. Within-subject factors were Load (low, high) and Hemisphere (contralateral parietal, ipsilateral parietal). Further analyses were undertaken on frontal theta activity and contralateral parietal alpha activity with a series of paired-samples *t*-tests on every time point (length between time points = 32 ms) between low and high load. Multiple comparisons were controlled for via the Benjamini & Hochberg (1995) procedure.

## 3.4 Results

### 3.4.1 Behavioural performance

To begin, behavioural performance values were analysed in both the visual search and tracking task with paired-samples *t*-tests. For visual search, accuracy and reaction time values were analysed and compared between visual search low load and visual search high load. For the pursuit tracking task which lasted approximately five seconds, two measures of immediate measures of task-switch readiness: time to first mouse movement, time to move the mouse inside the target, and two continuous measures from throughout the tracking task: the number of deviations away from the target, and the total duration of deviations were analysed. Each measure was calculated and compared immediately following visual search low load, and immediately following visual search high load.

#### *3.4.1.1 Visual search performance*

A paired-samples *t*-test revealed a significant difference in accuracy between the low load and high load conditions, $t_{(35)} = 7.95$, $p < .001$, $d_z = 1.34$. As expected, accuracy was greater during low load. Similarly, there was a significant difference in reaction time between low load and high load, $t_{(35)} = -18.33$, $p < 0.001$, $d_z = 3.06$. As expected, reaction time was greater, reflecting longer response times, during visual search high load. Overall, reaction time and accuracy measures reveal worse performance during visual search high load when compared to low load. See Figure 3.3 and Table 3.1 for means and SDs.

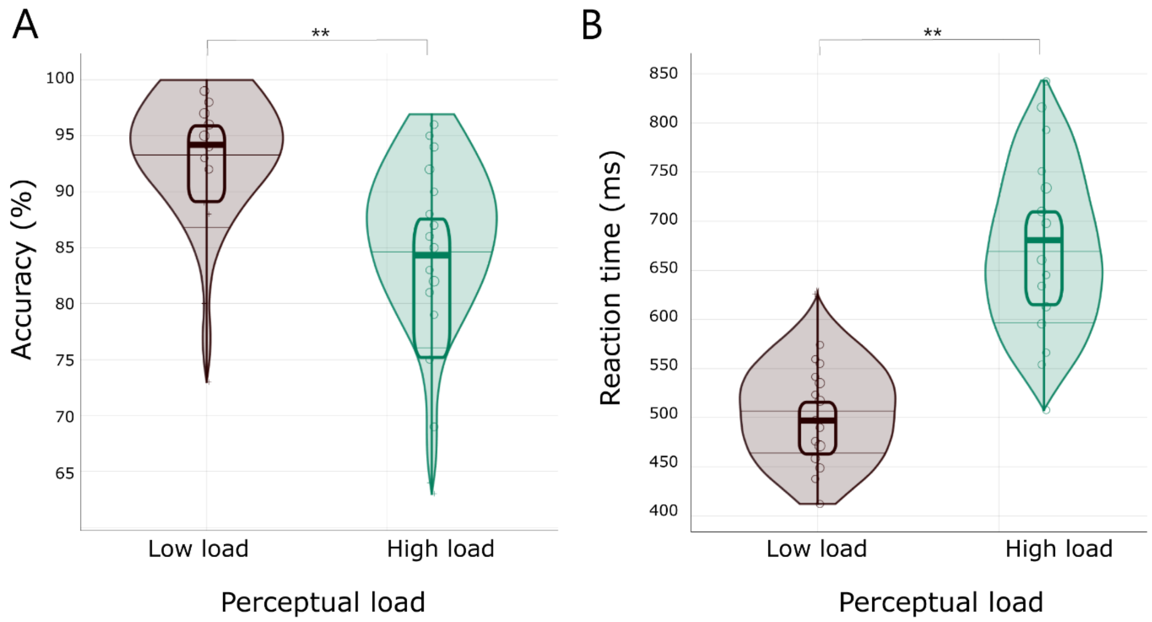



*Figure 3.3*. Visual search task performance data over low and high perceptual load. **(A)** Mean accuracy score (%). **(B)** Mean reaction time (ms). Violin plots represent the distribution of each data series, with a box plot and whisker drawn over these data (notch centre = mean). Key: ** represents $p$ < .001.

### *3.4.1.2 Tracking task performance*

Paired-samples *t*-tests revealed no significant differences between low and high load for time to first mouse movement, $t_{(35)} = -1.98$, $p = .06$, time to target, $t_{(35)} = 1.14$, $p = .26$, or frequency of deviations, $t_{(35)} = .61$, $p = .51$. A significant difference was found between low and high load for duration of deviations from the target, $t_{(35)} = -2.68$, $p = .01$, $d_z = -0.45$. On average, participants spent longer outside the target in the high load condition, compared to the low load condition. Overall, the results suggest that reaction time measures were not impacted by the attentional load of the prior visual search task, but participants continuous accuracy did suffer following high load compared to following low load. See Figure 3.4 and Table 3.1 for means and SDs.

*Table 3.1.* Mean (SD) of visual search and tracking task behavioural performance.

| Task | Measure | Low load | High load |
|---|---|---|---|
| Visual search | Accuracy (%)** | 94 (5) | 84 (9) |
| | Reaction time (ms)** | 496 (46) | 638 (80) |
| Tracking task | Time to first mouse movement (ms) | 696 (130) | 708 (123) |
| | Time to target (ms) | 524 (797) | 512 (744) |
| | Frequency of deviations | 5 (1) | 5 (1) |
| | Duration of deviations (ms)* | 148 (25) | 156 (27) |

*Key:* * represents $p \leq .01$; ** represents $p < .001$

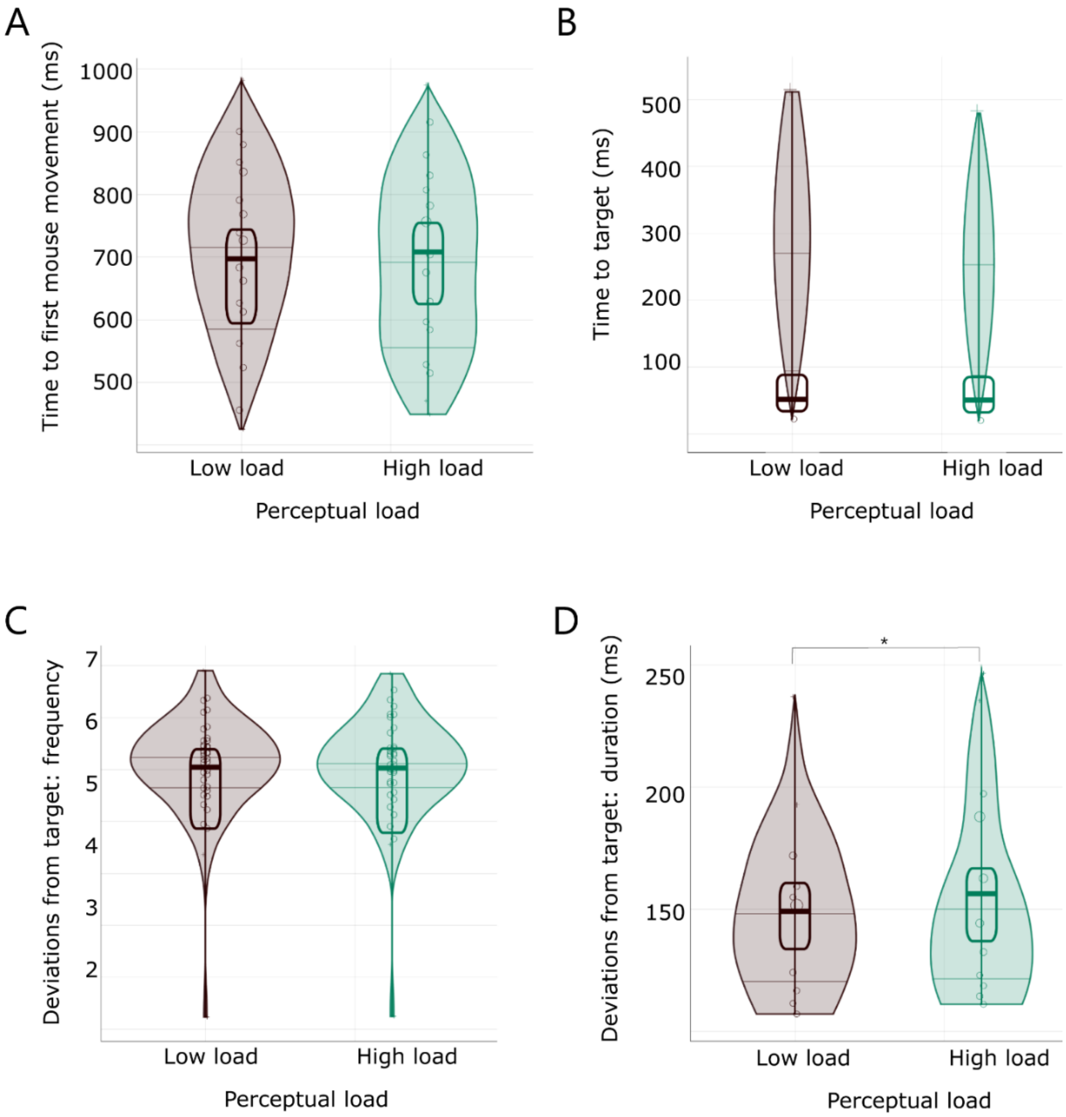



*Figure 3.4.* Tracking task performance data preceded by low and high perceptual load. **(A)** Mean time to first mouse grab (ms) **(B)** Mean time to target (ms) **(C)** Mean number of deviations from the target **(D)** Mean duration of deviations (ms). Key: * represents $p < .01$.

### 3.4.2 Visual search EEG

Next, EEG data from the visual search task was analysed. First, event-related potential (ERP) components the N1 and P3b were compared with *t*-tests between low and high visual search load. Then, spectral analyses were undertaken to explore alpha and theta dynamics between low and high visual search load. A series of *t*-tests were undertaken to explore frontal alpha power, frontal theta power, parieto-occipital alpha power, and parieto-occipital theta power.

#### *3.4.2.1 Event-related potentials (ERPs)*

A paired-samples *t*-test revealed that N1 amplitude did not statistically significantly differ between low and high load, $t_{(35)} = -0.73$, $p = .47$. N1 amplitude was similar during low load ($M = -5.97$, $SD = 3.74$) and high load ($M = -5.85$, $SD = 3.77$). A paired-samples *t*-test revealed P3b amplitude significantly differed between low and high load, $t_{(35)} = 7.64$, $p < .001$, $d_z = 1.27$. P3b amplitude was greater during low load ($M = 4.07$, $SD = 2.77$) when compared to high load ($M = 2.56$, $SD = 2.17$). See Figure 3.5.

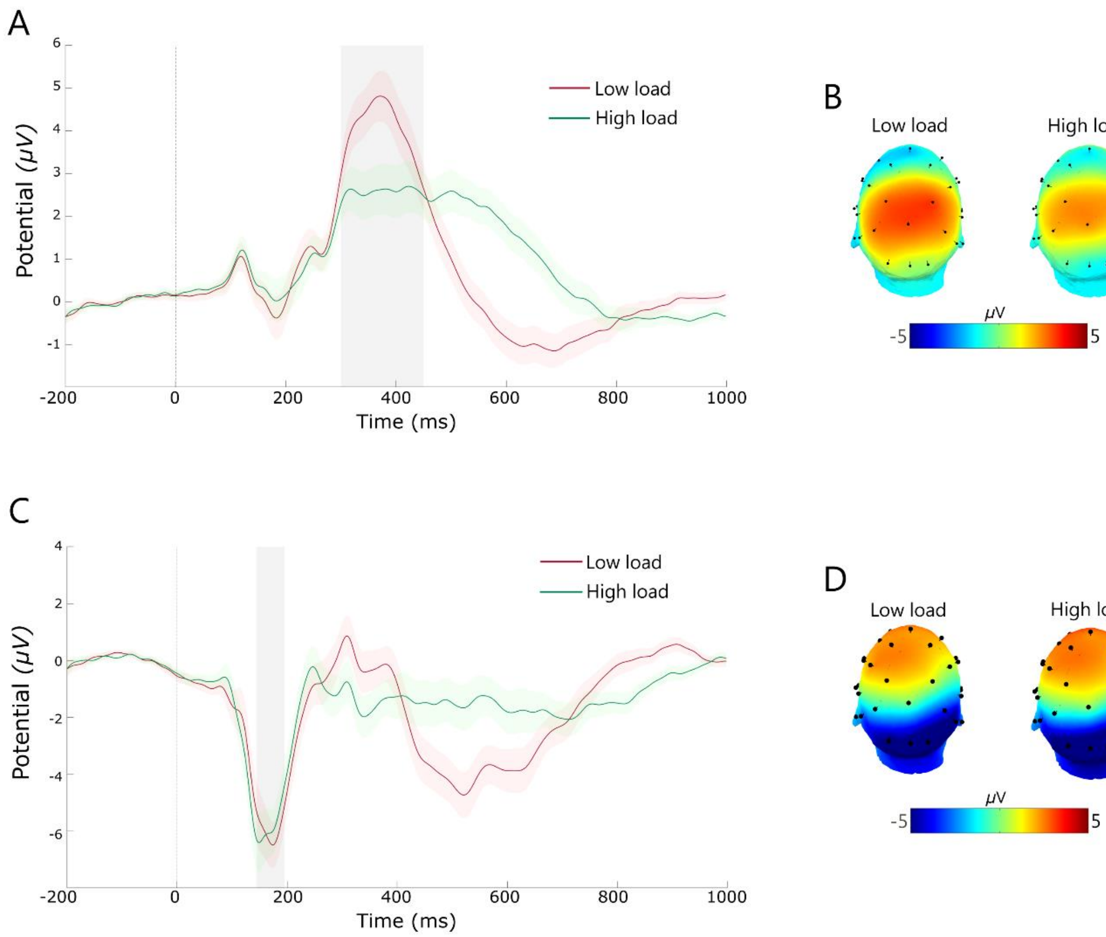



*Figure 3.5.* Grand-averaged P3b and N1 ERPs time-locked to visual search presentation under low and high perceptual load. Shaded areas represent the time course extracted for statistical analyses. Shaded error bars represent the ± standard error of the mean difference. **(A)** Grand-averaged P3b. Correct trials were averaged across midline parietal electrode Pz. **(B)** P3b (300 – 450ms) mean amplitude scalp topography for low and high perceptual load. **(C)** Grand-averaged N1. Correct trials were averaged across occipital sites: O1, O2, OZ. **(D)** N1 (145 – 195ms) mean amplitude scalp topography for low and high perceptual load.

### *3.4.2.2 Spectral analyses*

A paired-samples *t*-test revealed frontal theta ERS was greater during low load, $t_{(35)} = 2.92$, $p = .006$, $d_z = 0.49$, however a separate paired-samples *t*-test revealed parieto-occipital theta ERS did not differ between low and high load, $t_{(35)} = -0.08$, $p = .45$. A paired-samples *t*-test revealed frontal alpha ERD was significantly greater during high compared to low load, $t_{(35)} = 3.35$, $p = .002$, $d_z = 0.56$. Similarly, a paired-samples *t*-test revealed parieto-occipital alpha ERD was significantly greater during high compared to low load, $t_{(35)} = 3.00$, $p = .005$, $d_z = 0.50$. See Table 3.2 for means and SDs and Figure 3.6 for time-frequency spectrograms.

*Table 3.2.* Mean (SD) of visual search time-frequency power (dB normalised).

| Frequency (Hz) | Averaged electrodes | Low load | High load |
|---|---|---|---|
| Theta (4 – 7)* | Fz, FC1, FC2 | 1.02 (0.88) | 0.76 (0.74) |
| Theta (4 – 7) | P3, P7, P4, P8, O1, O2 | 0.89 (1.04) | 0.97 (1.08) |
| Alpha (8.5 – 12.5)* | Fz, FC1, FC2 | -1.00 (1.33) | -1.31 (1.45) |
| Alpha (8.5 – 12.5)* | P3, P7, P4, P8, O1, O2 | -2.19 (1.66) | -2.50 (1.88) |

*Key:* * represents $p$ < .01

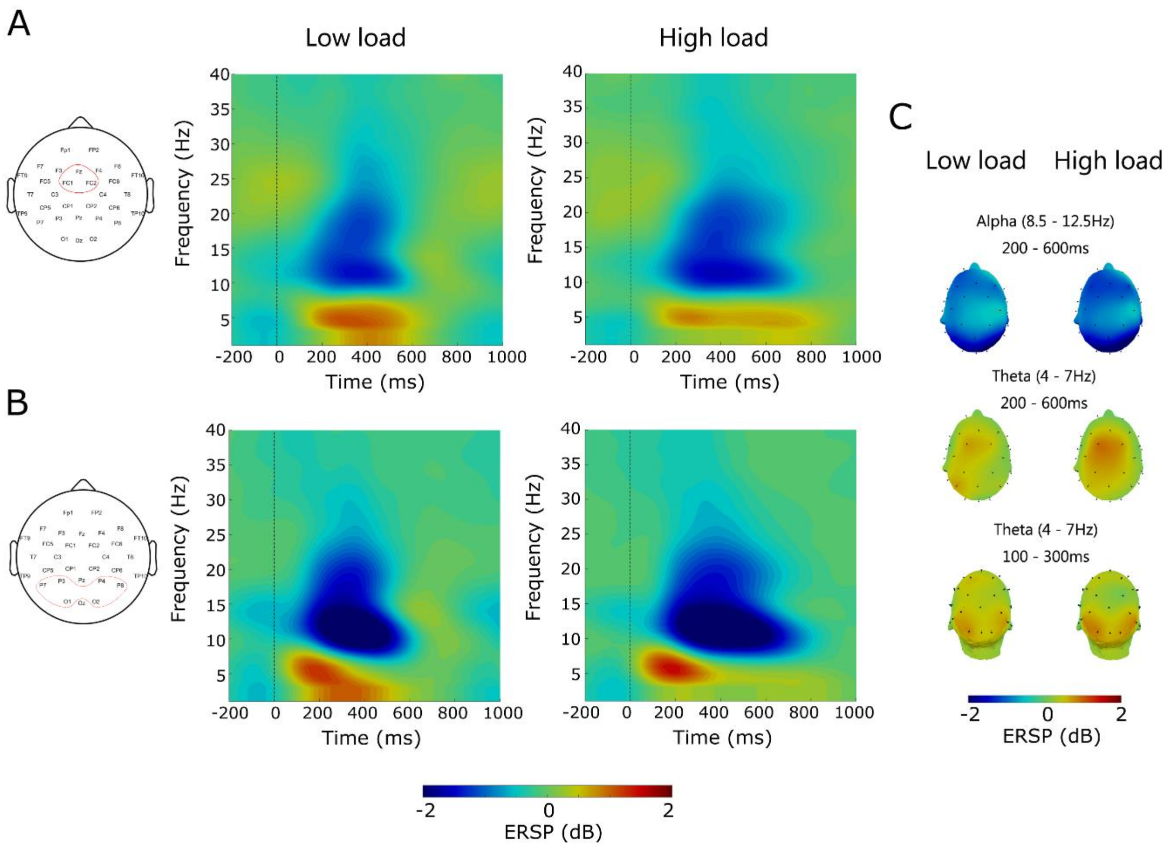



*Figure 3.6.* Time-frequency results for visual search activity over low and high perceptual load. **(A)** Grand-averaged time-frequency spectrograms for frontal (FC1, FC2, Fz) visual search activity. **(B)** Grand-averaged time-frequency spectrograms for parieto-occipital (P4, P7, P8, P3, O1, O2) visual search activity. **(C)** Averaged topographical plots over time points and frequencies of interest.

### 3.4.3 Tracking task EEG

Finally, EEG activity was analysed over the continuous 5 s visuomotor tracking task, temporally distinct from the visual search paradigm which consisted of a single button press. Analyses were interested in the impact of the attentional load of the preceding visual search task, and so the data compared were: tracking task following low load search, and tracking task following high load search. Similarly to visual search analysis, frontal theta, parietal theta, frontal alpha, parietal alpha, were analysed. Theta activity was a transient response apparent in the first second of the task and was therefore associated with the task switch. Alpha was apparent over the task period, lasting approximately 5 s, representing the temporal dynamics of attentional demand. As lateralisation effects were expected for attention-related parietal activity, parietal theta and parietal alpha were analysed with a 2 (Load: low, high) x 2 (Hemisphere: contralateral, ipsilateral) repeated measures ANOVA, while frontal theta and frontal alpha were analysed with paired samples *t*-tests.

#### *3.4.3.1 Theta activity*

A paired-samples *t*-test revealed frontal theta synchronisation did not statistically differ between low and high load, $t_{(35)} = 1.66$, $p = .10$. See Table 3.3 for an overview of the means and standard deviations. Additional analyses of paired-samples *t*-tests on each time point revealed that frontal theta power appeared later during high load. Theta power was significantly greater during low load at time points 110ms, 142ms, 172ms; whereas theta power was significantly greater during high load at time points 268ms, 300ms, 330ms, 362ms and 394ms. See Figure 3.7 for visualisation of results and Appendix 3.3 for *t*-test results.

A two-way repeated measures ANOVA was performed to understand the impact of low and high perceptual load on parietal theta activity during the tracking task. The ANOVA revealed no significant main effects for Load, $F_{(1,35)} = 0.93$, $p = .34$, nor Hemisphere, $F_{(1,35)} = 1.10$, $p = .30$. However, the interaction effect between Load and Hemi was significant, $F_{(1, 35)} = 5.70$, $p = .022$, $\eta_p^2 = 0.14$. However, once multiple comparisons were accounted for, the difference between the low load and high load condition in the contralateral hemisphere (P3, P7) were not significant ($p = .061$). See Table 3.3 for an overview of the means and standard deviations.

*Table 3.3.* Mean (SD) of tracking task theta (4 – 8 Hz) power (dB normalised).

| Electrodes | Time window (ms) | Low load | High load |
|---|---|---|---|
| Fz, FC1, FC2 | 100 – 400 | 2.73 (1.34) | 2.53 (1.37) |
| P3, P7 | 150 – 450 | 1.47 (1.39) | 1.21 (1.35) |
| P4, P8 | 150 – 450 | 1.57 (1.29) | 1.67 (1.14) |

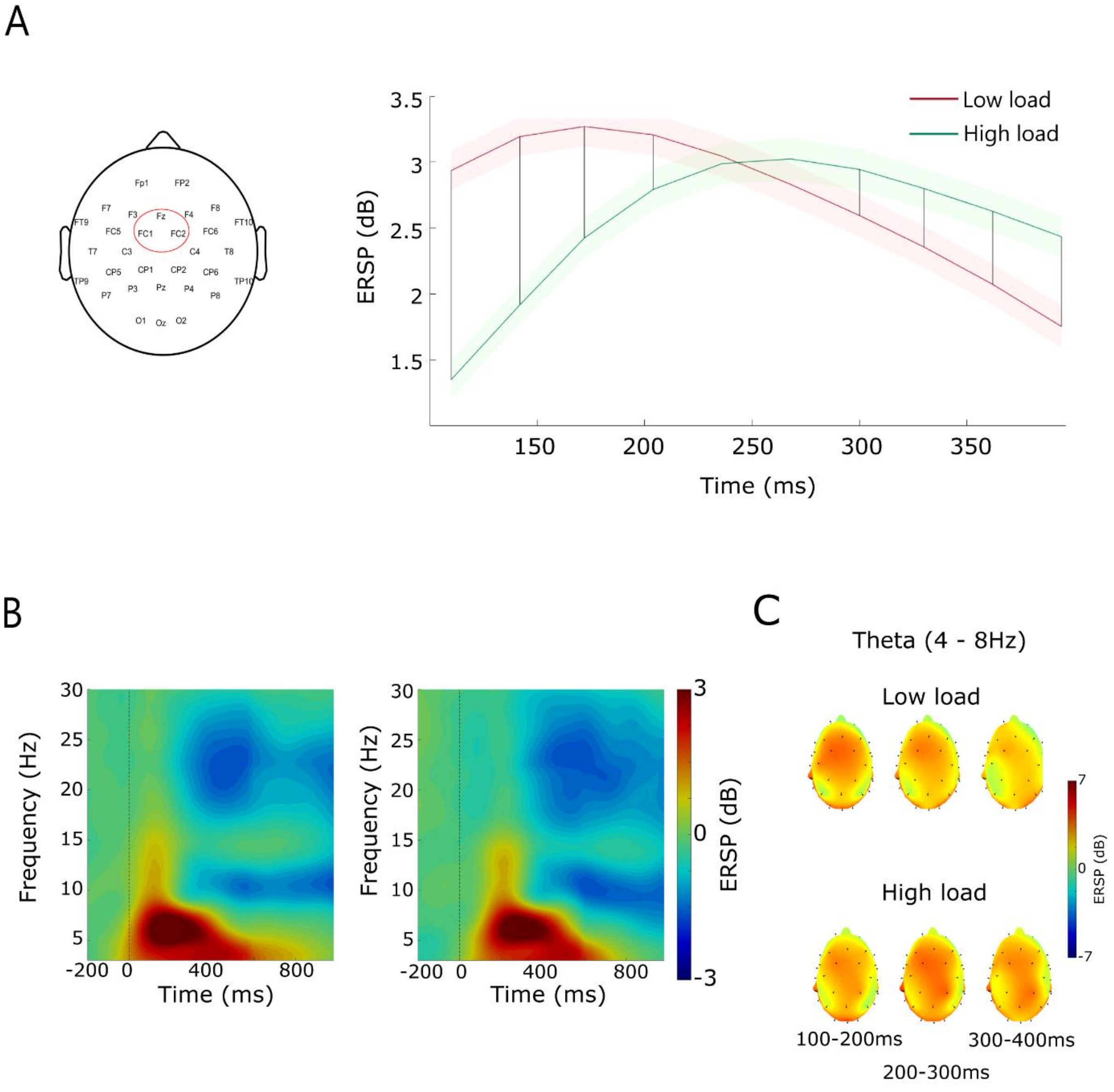



*Figure 3.7.* Grand-averaged dB normalised signal of frontal (FC1, FC2, Fz) theta (4 – 8 Hz) during the tracking task preceded by low or high perceptual load. **(A)** Line graph depicting evoked frontal theta activity over 100 – 400 ms post cue. Significant time points are depicted by a grey line ($p \leq .05$ corrected). Shaded areas represent the ± standard error of the mean difference. **(B)** Grand-averaged time-frequency spectrogram for frontal theta activity during low and high load. **(C)** Topographical difference (high – low load) plots represent theta activity averaged over three time series: 100 – 200 ms, 200 – 300 ms, 300 – 400 ms.

*3.4.3.2 Alpha activity*

A paired-samples *t*-test revealed frontal alpha desynchronisation did not statistically differ between low and high load, $t_{(35)} = 0.17$, $p = .97$. A two-way repeated measures ANOVA was performed to understand the impact of low and high perceptual load on parietal alpha activity during the tracking task. The ANOVA revealed an interaction effect between Load and Hemi, $F_{(1,35)} = 6.72$, $p = .01$, $\eta_p^2 = 0.16$. The main effect of Load was not significant, $F_{(1,35)} = 0.42$, $p = .52$, nor was the main effect of Hemisphere, $F_{(1,35)} = 0.06$, $p = .80$. *Post hoc* comparisons revealed that for the contralateral parietal hemisphere only (P3, P7), alpha desynchronisation was greater during high load compared to low load ($p = .05$). See Table 3.4 for an overview of the means and SDs. Additional analyses of paired-samples *t*-tests on each time point revealed contralateral parietal alpha desynchronisation was greater during high load at time points 268ms, 552ms, 584ms, 646 – 710ms, 3868 – 4056ms, 4310 – 4436ms, 4594 – 4942ms, and 5004 – 5098ms after cue onset (see Figure 3.8 for a graphical representation of the results and Appendix 3.4 for *t*-test results).

*Table 3.4.* Mean (SD) of tracking task alpha (8.5 – 12.5 Hz) power (dB normalised).

| Electrodes | Time window (ms) | Low load | High load |
|---|---|---|---|
| Fz, FC1, FC2 | 350 – 5100 | -1.26 (1.49) | -1.26 (1.27) |
| P3, P7* | 250 – 5100 | -1.91 (1.70) | -2.16 (1.77) |
| P4, P8 | 250 – 5100 | -2.13 (1.92) | -2.03 (1.91) |

*Key:* * represents $p \leq .05$

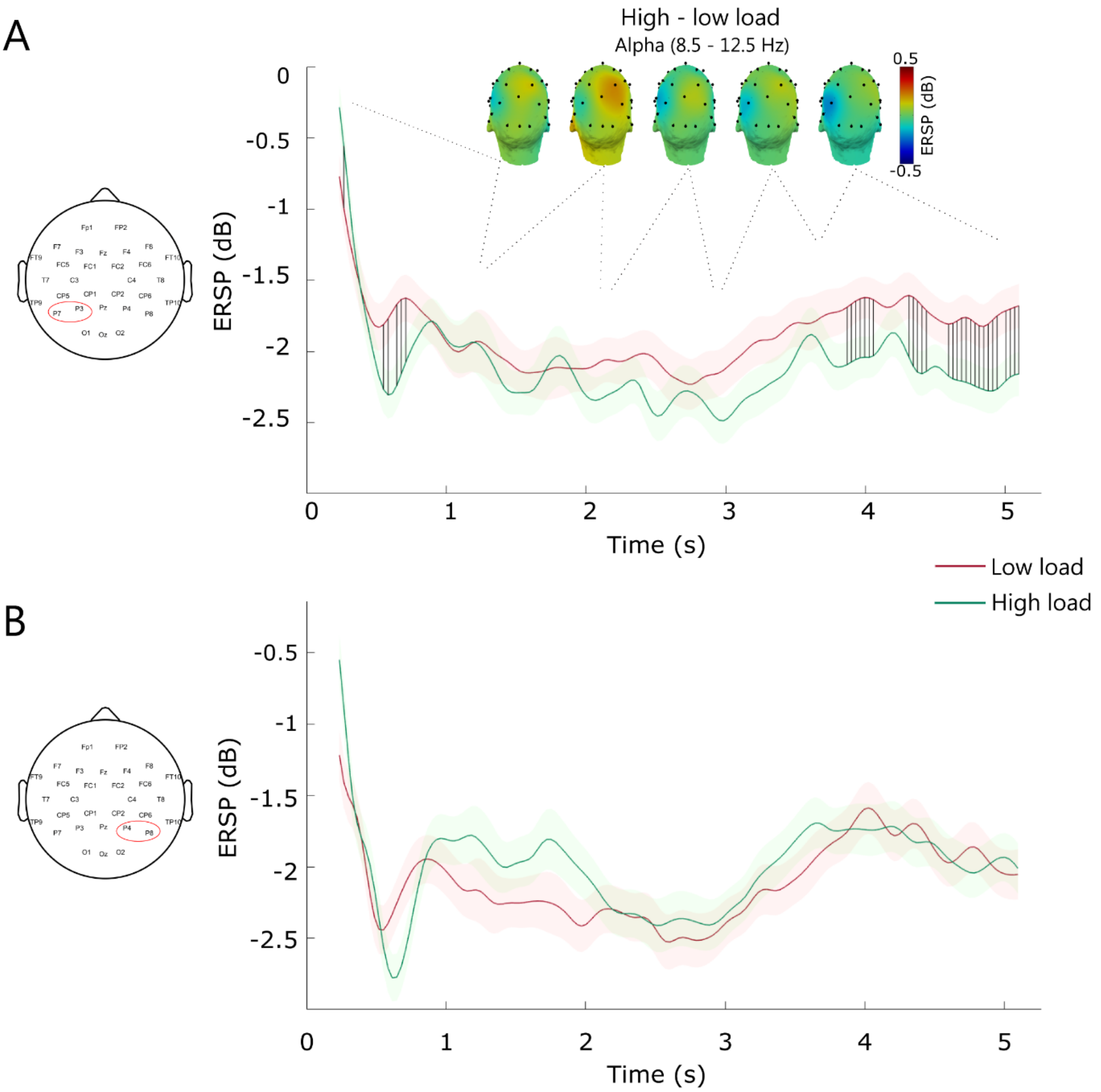



*Figure 3.8.* Grand-averaged dB normalised signal of parietal alpha (8.5 – 12.5 Hz) during the tracking task preceded by low or high perceptual load. Significant time points are depicted by a grey line ($p \leq .05$ corrected). Shaded areas represent the ± standard error of the mean difference. **(A)** Contralateral parietal (averaged P3, P7) alpha activity over low and high load. Topographical difference (high – low load) plots represent alpha activity averaged over five time series: 250 – 1220 ms, 1220 – 2190 ms, 2190 – 3160 ms, 3160 – 4130 ms, 4130 – 5100 ms. **(B)** Ipsilateral parietal (averaged P4, P8) alpha activity over low and high load.

*3.4.3.3 Motor activity*

Control analyses were run to ensure that hemispheric alpha differences were not related to motor activity, as all participants used their right hand to respond to the tracking task. Paired samples *t*-test revealed no significant differences in contralateral (C3) motor beta activity between low (*M* = -2.09, *SD* = 1.55) and high load (*M* = -2.08, *SD* = 1.79), $t_{(35)} = -0.09$, $p = .93$; nor was there a statistically significant difference in alpha activity between low (*M* = -1.88, *SD* = 1.69) and high load (*M* = -1.9, *SD* = 1.78), $t_{(35)} = 0.23$, $p = .82$. See Appendix 3.5 for condition grand-averaged time-frequency spectrogram for C3 representing motor-related oscillatory activity.

### 3.4.4 Overview of results

An overview of visual search and tracking task results are presented in Table 3.5. In summary, performance was worse during the high load visual search task when compared to the low load visual search task, as represented by reduced accuracy and increased reaction time. The N1 component did not differ between low or high perceptual load, however, the P3b, indicating target detection, was greater during low load. Theta ERS was greater during low load, and frontal and parieto-occipital alpha activity was lower (greater alpha ERD) during high load. The continuous tracking task revealed that participants spent longer outside the tracking target following high load, however reaction time measures (i.e. time to first mouse movement, time to target) and the number of deviations did not differ. Transient frontal ERS was similar between conditions but was significantly delayed following high load. Contralateral parietal theta ERS was lower following high load compared to low load, but this did not survive Bonferroni corrections. Frontal alpha ERD did not differ following low and high load. Contralateral, but not ipsilateral, parietal alpha activity was lower (greater alpha ERD) following high load. Motor-related activity did not differ following low and high load.

*Table 3.5.* Overview of results during visual search and pursuit tracking.

| Measure | Parameter | Time window (ms) | High compared to low load |
|---|---|---|---|
| **Visual search** | | | |
| Behavioural | Accuracy (%) | – | ↓ |
| | Reaction time (ms) | – | ↑ |
| EEG | N1 (µV) | 145 – 195 | - |
| | P3b (µV) | 300 – 450 | ↓ |
| | Frontal theta ERS (dB) | 200 – 600 | ↓ |
| | Frontal alpha ERD (dB) | 200 – 600 | ↑ |
| | Parieto-occipital theta ERS (dB) | 100 – 300 | – |
| | Parieto-occipital alpha ERD (dB) | 200 – 600 | ↑ |
| **Tracking task** | | | |
| Behavioural | Time to first mouse movement (ms) | – | – |
| | Time to target (ms) | – | – |
| | Frequency of deviations | – | – |
| | Duration of deviations (ms) | – | ↑ |
| EEG | Frontal alpha ERD (dB) | 350 – 5100 | – |
| | Frontal theta ERS (dB) | 100 – 400 | Delayed |
| | Contralateral parietal alpha ERD (dB) | 250 – 5100 | ↑ |
| | Ipsilateral parietal alpha ERD (dB) | 250 – 5100 | – |
| | Contralateral parietal theta ERS (dB) | 150 – 450 | ↓* |
| | Ipsilateral parietal theta ERS (dB) | 150 – 450 | – |
| | Contralateral motor beta ERD (dB) | 150 – 5100 | – |
| | Contralateral motor alpha ERD (dB) | 250 – 5100 | – |

*Key:* '↑' represents the parameter was significantly greater during high load, '↓' was significantly lower during high load, ↓* represents the parameter was significantly lower during high load, but the difference did not survive Bonferroni corrections, and '–' represents the parameter was similar between low and high load. Key: event-related synchronisation: ERS; event-related desynchronisation: ERD; dB: decibel normalised

## 3.5 Discussion

The aim of this study was to determine the temporal characteristics of neural oscillatory activity and behavioural performance during a visuomotor task preceded by a visual search task differing in attentional load. Behavioural, event-related potential (ERP), and spectral analyses of the visual search supported previous research indicating that a task of high perceptual load reduces accuracy and requires additional recruitment of target encoding and attentional mechanisms. Tracking task analyses identified three main findings. The first finding was in the performance data. Participants reaction time following cue presentation was similar between low and high load. However, continuous performance measures revealed that participants spent longer outside the tracking target, representing poorer accuracy. The second finding concerned frontal theta activity. Frontal theta synchronisation (ERS) was approximately 100 ms delayed during the tracking task when following high load, compared to low load. The final finding revealed sustained desynchronisation (ERD) during the tracking task. Alpha ERD was greatest in the contralateral (to the moving hand) parietal area following high load. Further analyses revealed that two time series represented the significant difference. The first time series started at approximately 0.5 until 0.7 s after cue onset. The second time series spanned the end of the 5 s tracking task, starting at 3.9 s.

### 3.5.1 Performance was worse, P3b amplitude was suppressed, alpha desynchronisation was pronounced, and frontal theta synchronisation was lower, during high load visual search

Posterior P3b amplitude was greater in the low perceptual load condition, compared to high perceptual load during visual search. This strong effect was expected and replicates earlier studies demonstrating smaller P3bs elicited by visual stimuli in tasks of high working memory load (Gevins & Cutillo, 1993), increased task difficulty (Polich, 1987), and high cognitive load (Giraudet et al., 2015). The reduced amplitude during high load suggests a suppressed ability to attend to target stimuli and fewer targets correctly identified. No differences in N1 amplitude between loads suggest that the differences in P3b and behavioural results are the result from cognitive rather than perceptual processes.

The P3b amplitude was paralleled by alpha power ERD in the parieto-occipital regions. During high load, alpha ERD was greater between 200 and 600 ms. ERD was apparent beyond the P3b time window, and therefore suggests that P3b and alpha differences reflect different mechanisms of the same underlying cognitive process. Together, the P3b and alpha desynchronisation results suggest a greater recruitment of attentional mechanisms during high load. Frontal alpha ERD differences were also found, supporting the notion that attention-related alpha is under control of a fronto-parietal network (Foxe & Snyder, 2011). Frontal alpha preceded parietal alpha, representing distinct top-down processes involved in a feed-forward mechanism during visual search.

Frontal theta event-related synchronisation (ERS) was stronger during visual search low load. These results were unexpected, as theta power is often positively related to task demands and demonstrates a linear relationship with task difficulty (e.g. Puma et al., 2018; Zakrzewska & Brzezicka, 2014). However, the present study findings are in line with the view that frontal theta synchronisation is related to movement selection during visuomotor processing (Rawle et al., 2012). As the behavioural responses indicate, participants were slower at selecting the correct target (either M or N) during high load, which may represent uncertainty of movement selection. Rawle et al. (2012) found that frontal theta power was lower following a visual search task with four potential targets compared to two targets. Therefore, movement uncertainty has previously been represented by lower frontal theta power.

No differences between load were found in parietal theta. Rawle et al. (2012) argued that parietal theta is associated with visuospatial representation (i.e. the spatial representation of visual objects), but not to the complexity of those objects. This explanation fits the present study as the spatial representation of the visual search objects were identical between conditions, as only complexity was modulated.

### 3.5.2 Reaction times did not differ, but accuracy suffered, during the tracking task following high load

Tracking time reaction time measures did not differ between prior perceptual low and high load conditions. However, participants spent longer outside the target following high load, reflecting poorer accuracy. There could be several reasons why reaction times did not differ between low and high tracking task conditions. Firstly, it is well demonstrated that the smaller the cue-stimulus interval (CSI), the greater the reaction time is to a task switch (Meiran, 1996). In this study, task cues were presented with a small CSI of 100ms. Task-set reconfiguration theory suggests that prolonged reaction time results from increased difficulty in preparing for the new task-set (Rogers & Monsell, 1995). The small CSI may instigate poorer instantaneous performance regardless of prior perceptual load.

Therefore, reaction times may have exhibited a floor effect. Comparing behavioural results with short and long CSIs (e.g. 600ms; Li et al., 2012) will be able to elucidate further.

Although participants were instructed to be as quick and as accurate as they could, individual differences will emerge between how much emphasis they give to speed or accuracy, known as the speed-accuracy trade off (Heitz, 2014). Therefore, participants who sacrificed accuracy for speed, will look as if they performed better in reaction time measures. In addition, cognitive processes, such as working memory, which are highly implicated in task switching, have been related to accuracy rather than reaction time (Unsworth & Engle, 2008). Therefore, the present paradigm facilitated capture of performance errors due to a cognitive cost throughout the task, which were not reflected via transient reaction time measures.

The continuous measures of performance found that participants spent longer outside the target during pursuit tracking, when following the visual search task of high perceptual load. The size of the effect (0.45) was medium (> 0.35), and is an important finding, as many tasks in the real world involve continuous efforts to maintain appropriate performance, instead of occasional impulsive behavioural choices (i.e. selective button presses). Accuracy during a pursuit tracking task has been shown to improve with feedback (Lang et al., 2011), visual enhancements (Park & Park, 2007), and predictability (Broeker et al., 2020). Therefore, the behavioural results suggest that pursuit tracking when following a task of high perceptual load may benefit from additional input that has been shown to improve performance. In addition, given that reward processing has been shown to impact sustained attention (see Chapter 1.1.2), future research manipulating feedback or reward during pursuit tracking following an attentional task of low or high load may provide further insight performance fluctuations during naturalistic sustained attention tasks. Would rewards associated with accurate tracking improve tracking performance regardless of the attentional load of the previous task?

### 3.5.3 Evoked frontal theta was delayed during the tracking task following high load

In contrast to alpha, frontal theta activity increased rapidly at cue onset, and lasted for a few hundred milliseconds. Although theta power was similar between load, additional analyses revealed that it was delayed by approximately 100 ms during high load. Still in line with the movement selection literature, frontal theta synchronisation has been associated with target encoding mechanisms (i.e. Herweg et al., 2020; Proskovec et al., 2018). Yet, it is often cited to have an unspecific function, and is related to cognitive control processes more generally (Cavanagh & Frank, 2014). As it was hypothesised that switching to a new task following high load would be more difficult than switching to a new task following low load, frontal theta activity sharing similar power values between conditions was unexpected. Yet, theta power is often related to instant measures of task performance. For example, Cooper et al. (2019) recently found that post-target theta power increased with faster reaction times during task switching. Therefore, the lack of differences between theta power values corresponds to the lack of differences between response time behavioural measures.

Yet, further analyses revealed that theta power was approximately 100ms delayed when following cue presentation after high load. As the cue represented an event that indicated a need for increased control, the frontal theta lag represents delayed active cortical processing mechanisms. It is important to note that this had no impact on instant measures of performance i.e. participants grabbed the mouse at similar times between loads. However, accuracy was impacted, with greater time spent outside the target following high load. Therefore, theta power may still be involved in the actual computations needed for task performance, however the theta lag did not have any impact on instant measures of performance.

### 3.5.4 Parietal alpha desynchronisation was greater during the tracking task following high load

The present study revealed a sustained alpha reduction, reflecting desynchronisation, in the contralateral parietal hemisphere (relative to hand movement) during the tracking task following high perceptual load. Although alpha lateralisation effects have been mainly attributed to visuospatial tasks (i.e. spatial cueing paradigm: Doesburg et al., 2016; Worden et al., 2000), research has demonstrated lateralisation effects in other systems such as somatosensory (Haegens et al., 2011). Earlier studies of visuomotor tasks have demonstrated stronger alpha ERD effects over contralateral parietal areas before and during movement execution (Babiloni et al., 1999). In addition, stronger alpha ERD effects have been found in contralateral areas during more complex visuomotor tasks (van der Meer & van der Weel, 2017). As alpha is important for functional engagement and disengagement of specific regions, greater contralateral alpha ERD supports the view of increased recruitment of visuomotor attentional mechanisms. Other studies have demonstrated a general effect for the left parietal cortex on motor attention (Mutha et al., 2013). For example, lesion studies have demonstrated deficits in movement selection (Castiello & Paine, 2002), and more recent work found that transcranial magnetic stimulation of the left parietal cortex delayed movement execution (Oliveira et al., 2010). Together, these results suggest a left hemisphere dominance for execution of movements.

Huang et al. (2007) found alpha increases over a contralateral left somatomotor cluster for high errors in a continuous tracking task. The authors suggest that this might reflect impairments in somatomotor attention (Bauer et al., 2006; Worden et al., 2000), supported by poorer performance and tonic increases in occipital power most likely reflecting drowsiness. The present study findings add to this research by demonstrating lower parietal alpha power and poorer accuracy during a task of high load. Potentially, this could represent a greater recruitment of visuomotor attention processes. However, considering the present study used right-handed participants only, it remains unclear whether this increase in visuomotor attention reflects a genuine difference between left and right hemispheres, or whether it is due to movement execution of the right hand linked to left hemisphere motor centres.

Alpha ERD started at approximately 250 ms after cue onset during the tracking task following both low and high load visual search. The disengagement of attention from the location of the previous task would be necessary before attention could be reoriented toward the new task location, and this would plausibly cause a delay in the engagement of top–down mechanisms. This delay was similar between conditions, which is supported by the behavioural data which found that response times were also similar between loads. Two significant time regions were identified representing the differences between low and high load. The first was at the beginning of the tracking task, up until 750 ms after cue onset, most likely reflecting evoked alpha activity relating to the task switch. Similar to previous research, Proskovec et al. (2019) found parietal alpha ERD starting at approximately 350 ms after stimulus onset. Using a classic cue paradigm, alpha desynchronisation persisted during the task period, up until 1550 ms. Applying beamformer analysis, they found that alpha activity in the left (contra) intraparietal sulcus and superior parietal lobule were sustained from 350 to 750 ms and 750 to 1150 ms (after movement). These areas which are part of the dorsal attention network have been implemented in top-down allocation of attentional control (Behrmann et al., 2004), and therefore implies that sustained alpha ERD may play a critical role in attentional processes. This may be in part due to interference from the previous task. The second time region of interest was from 3.9 s until the end of the task. This demonstrates a lasting attentional cost, supporting the notion that a task switch is impacted by the difficulty of a previous task (i.e. the visual search task). Why was alpha ERD not suppressed throughout the whole tracking task? This second time region of alpha modulation may represent a pronounced requirement for neural activity modulating attentional mechanisms. The tracking task required 5 seconds of continual effort. At the task switch, alpha ERD rapidly increased, representing the initial selective enhancement of neural activity responsible for processing stimuli. Following low load, attention-related alpha recovered at a quicker rate compared to high load. This is represented by the significant differences found between load during the end of the tracking task. Therefore, at around 4 s, recruitment of attentional mechanisms was pronounced during high load and likely due to the continual need for selective enhancement of neural responsiveness to modulate the engagement of task-relevant neurons and the disengagement of neurons responsible for processing irrelevant aspects related to the tracking task (Foxe & Snyder, 2011). The behavioural results assist with this interpretation. Lower parietal alpha ERD (i.e. greater alpha power) is often related to poorer performance (Clayton et al., 2015). During high load, alpha ERD may have increased during the first 750 ms to support task performance, indicated by similar reaction time measures during low and high load. Although accuracy measures indicated poorer performance; these measures were averaged over the complete tracking task period (5 s). Therefore, it is possible that participants were less accurate when alpha ERD reduced (i.e. alpha increased) during high load. During this period, alpha power was similar to low load, suggesting that attentional mechanisms were impaired which may have resulted in poorer performance. However, further research is needed to confirm this interpretation. Measuring alpha ERD before and after tracking task errors (i.e. when the cursor drifts away from the tracking target) would improve understanding of the relationship between task performance and alpha ERD during continuous tasks.

### 3.5.5 Limitations

EEG signals are easily contaminated by artefacts of physiological and non-physiological origin. The challenge of physiological artefacts is particularly relevant for visuomotor tasks, which not only induce muscle-related artefacts, but also ocular artefacts due to eyelid movements and movements of the retinal dipoles (Croft & Barry, 2000), and both muscle and ocular signals are relatively strong in comparison to EEG. Muscular artefacts are characterised as high-frequency activity, and surface electromyography has shown that its frequency characteristics range from 20 – 600 Hz (Criswell, 2011). Therefore, in motor tasks, interpretation of data over 20 Hz is limited due to the impact of muscle activity increasing high-frequency power (Muthukumaraswamy, 2013). Therefore, the present study focused on the temporal evolution of low-frequency oscillatory activity, namely theta ERS and alpha ERD. Frequency bands such as gamma (30+ Hz) were not considered as broadband gamma was contaminated by electromyographic activity. In addition, the tracking task did not differ between conditions, minimising the impact of differential artefacts.

Unlike the majority of cognitive neuroscience experiments that aim to isolate processes, the current EEG study utilised a tracking task which required smooth-pursuit eye-movements. Consequently, it becomes unclear whether the neural mechanisms are related to ocular control or cognitive behaviours. There is a growing debate in the literature whether covert and overt shifts of attention are independent (Posner & Petersen, 1990), interdependent (Corbetta, 1998), or dependent to one another (Rizzolatti et al., 1987). However, many studies have found that neural activation is similar between covert and overt shifts of attention, and argue that covert shifts are simply an unexecuted movement, suggesting only ocular motor components differ (Beauchamp et al., 2001; de Haan et al., 2008). Similarly, the theta and alpha patterns observed were similar to the patterns found in the task switching and attention reorientation literature which manipulate overt shifts of attention (i.e. Proskovec et al., 2018, 2019). However, this is not to say that ocular movements did not impact these low-frequency mechanisms. The frontal eye fields play a critical role in the allocation of spatial attention and have been shown to interact with parietal alpha (Capotosto et al., 2009). However, it is important to adopt this approach to improve the transferability of the results to applied contexts, which include interactions between ocular and cognitive behaviours under naturalistic conditions.

One of the main concerns of measuring attention-related alpha activity during a tracking task is whether it can be separated from motor-related alpha activity, commonly known as mu. Like parietal alpha activity, mu is suppressed during task engagement (Pfurtscheller et al., 1997). Therefore, during a motor task, mu is suppressed over the sensorimotor cortices. Often, it is argued that it is not clear whether differences in alpha reductions are related to motor or cognitive components of the task (Anderson & Ding, 2011). However, alpha and mu are distinguishable via their topography. Mu arises specifically from the sensorimotor area, and therefore is measurable via EEG electrodes over the motor cortex. Because of this potential confound, control analyses for motor-related alpha activity were undertaken by contrasting motor cortex activity between low and high load (see Chapter 3.4.3.3). Results found there were no significant differences in motor-related oscillatory activity, suggesting parietal alpha activity was distinguishable to motor components.

### 3.5.6 Conclusion

Whereas previous studies have examined the specific effects of higher cognitive processes, this study explored prior attentional load on the general mechanisms of visuomotor performance. To this end, prior perceptual load (low versus high) was altered while keeping the physical properties of the secondary task - visuomotor pursuit tracking - the same. As supported by the literature, behavioural and EEG findings of the visual search task revealed perceptual load was successfully manipulated as demonstrated by increased errors and reaction times, increased movement uncertainty (frontal theta

ERS), and increased recruitment of attentional mechanisms (parieto-occipital alpha ERD) during high load. During the tracking task, short-lasting bursts of theta ERS and sustained alpha ERD were most pronounced at frontal and parietal scalp locations, respectively. This convergence supports the recruitment of a fronto-parietal network similar to the task switching and attentional re-orientation literature. Although theta power representing stimuli encoding neural mechanisms and behavioural response times did not differ between load, attention-related alpha activity was more pronounced at the beginning and end of the tracking task following high load. Participants' accuracy also suffered following high load, deviating from the target more so when compared to low load.

This study provides important implications for naturalistic, safety-critical, sustained attention activities. It demonstrates that the attentional demands of a prior task have detrimental effects on a sustained visuomotor task. If sudden intervention is required, such as switching out of autopilot during flight, or taking over in a semi-autonomous vehicle, then the characteristics of a prior task may have a significant impact on a safe and timely behavioural response. In addition, it provides insight into the performance decrements related to attentional and cognitive dysfunction in attention-related clinical disorders. Consequently, this study provides support for neurophysiological monitoring during human-machine interactions and to improve and assist clinical interventions. However, this study utilised a computer-based tracking task featuring elementary, two-dimensional, geometric shapes, as a surrogate for a real-world visuomotor task. In addition, a computer-based tracking task engages effortful controlled processes whereas real-world tasks such as driving will engage both effortful controlled processes and effortless automatic processes in experienced drivers. Therefore, it is not clear whether the same neural mechanisms will still be impacted during a more ecologically valid task. The next study described in Chapter 4 will explore whether these results are similar for a semi-naturalistic task, simulated driving, which will provide further insight into the possibility for utilising a biofeedback system, such as a driver state monitoring system, to improve the safety and wellbeing during semi-autonomous driving.

# 4.0 The Effect of Prior Attentional Load on Task Performance and Neural Oscillations During Simulated Driving

## 4.1 Overview

To improve the transferability of Chapters 3 findings towards semi-autonomous driving and to investigate the potential effectiveness of a driver state monitoring system (DSM), this study investigated whether event-related frontal and parieto-occipital alpha and theta oscillatory mechanisms were implicated during simulated driving when preceded by a task of differing attentional load. Participants partook in a study that required them to switch between an unrelated task and simulated driving. Three unrelated tasks were administered: a visual search task of low perceptual load, a visual search task of high perceptual load, and a passive viewing task involving monitoring the driving environment which acted as a control task. Reaction time, lateral and longitudinal driving control, as well as alpha and theta activity, were measured continuously throughout.

## 4.2 Introduction

One of the assumptions of Level 2 to 4 autonomous vehicle systems is that humans will take control of the vehicle. This process is known as the takeover process, and transitions can occur in planned or unplanned ways. For example, the vehicle always hands over control when at a roundabout, or during a system failure. System failures can, and will occur, for example, if the vehicle does not detect environmental changes due to changes in wind speed. Takeovers may therefore happen in emergency situations, where the autonomous vehicle system has failed to notice a difficult precipitating event and does not send a preceding takeover request. The driver must act quickly and safely to regain control of the vehicle. Yet, to safely regain control of the vehicle, depends on the extent to which the driver has remained engaged with the driving environment and their physical readiness (Zeeb et al., 2015).

### 4.2.1 Takeover time and takeover quality

Temporal measures used to assess takeover performance fall under two categories: takeover time and takeover quality. Takeover time is considered an instantaneous reaction time measure of transition, whereas takeover quality is a continuous measure of driving performance. Takeover time is generally considered the time elapsed from the takeover request and a driving behaviour from the human, such as a steering or pedal input. It is therefore an immediate measure of performance. Motor behaviour is commonly measured, such as when the driver activates the indicator (e.g. Li et al., 2018) or when the driver steers or brakes at certain thresholds (e.g. Feldhutter et al., 2018). For example, Gold et al. (2013) considered a takeover successful once the driver turned the steering wheel more than 2 degrees and altered the braking pedal position by 10%. In addition, gaze reaction time metrics are commonly assessed, including gaze to the forward roadway (e.g. Zeeb et al., 2016), gaze to the side mirror, and gaze to the speedometer (e.g. Vogelpohl et al., 2018).

It is acknowledged that a short takeover time does not necessarily indicate the passive driver has increased situation awareness nor may they be engaged with the driving environment. For example, Dogan et al. (2019) found that during takeover, a small number of drivers overtook on the inside which was considered an aberrant driving behaviour. In addition, aggressive braking could hint that drivers have reduced situation awareness. Therefore, measures of continuous performance are important to infer accuracy. Takeover quality is considered the driving performance after the takeover has successfully occurred. A number of takeover quality metrics have been used in the literature, including lateral acceleration (Körber et al., 2018), longitudinal acceleration (Feldhutter et al., 2018), lane position (Zeeb et al., 2017), and minimum distance headway to the lead vehicle (Madigan et al., 2018). Variability, in terms of standard deviation, are also utilised, the most commonly being the standard deviation of lateral position (SDLP; Merat et al., 2014). Behavioural studies have consistently demonstrated the negative impact of periods of automated driving on takeover performance. Gold et al. (2013) demonstrated that a cue five to seven seconds in advance was required for avoidance of a stationary object. Zhang et al.'s (2019) meta-analysis revealed an average takeover time of 2.72 seconds, though often time budgets between five to seven seconds are applied. Yet, takeover quality metrics provide evidence that the driving task is adversely affected for an extensive period of time. For example, Merat et al. (2014) found that participants took 35 to 40 seconds to stabilise their lateral control of a driving simulator following a period of automated driving. Several factors have been shown to modulate driving performance during partially automated driving. A recent meta-analysis revealed that takeover time was substantially impacted by factors such as lack of notifications, lack of experience, criticality of the situation, and performing secondary tasks (Zhang et al., 2019).

### 4.2.2 Non-driving related tasks and driving performance

Automated driving systems will alter drivers' behaviour and encourage involvement with non-driving related tasks (Carsten et al., 2012). Research with partially automated systems has already demonstrated that participants are more likely to shift their attention towards a secondary task, which in turn has a negative impact on driving performance (e.g. Rudin-Brown & Parker, 2004). Therefore, it is important to understand the relationship between secondary tasks and driving performance, and the mechanisms that drive this relationship. A growing body of literature has demonstrated the negative impact of these tasks on takeover performance. Dogan et al. (2019) tested participants undertaking two different secondary tasks: writing emails and watching videos. Participants had to retake control of the vehicle under two different situations depending on complexity: either obstacle avoidance or missing lane markings. Although they found that lateral and longitudinal control were negatively impacted when compared to manual driving in the missing lane markings condition, there was no effect of type of task on either critical situation. Moreover, participants did not subjectively rate any differences in workload between secondary tasks. These results are in line with a number of studies that have also found that the type of task does not differentially influence takeover performance (e.g. Zeeb et al., 2017; Bueno et al., 2016). Conversely, there are a handful of studies that have found an effect of type of task. Jarosch et al. (2019) investigated the effects of two tasks: a simple visual search task and a complex quiz. After 50 minutes of automated driving, a takeover was requested. Participants responded quicker in the quiz condition by braking faster and focusing on the road centre earlier. The authors suggest that the visual search task induced fatigue effects resulting in increased reaction times. The inconsistency between Jarosch et al.'s (2019) and Dogan et al.'s (2019) findings may be attributed to journey time and the duration of automated driving. In Dogan et al. (2019) the takeover request was requested after only 10-minutes of automated driving, unlikely to induce fatigue.

It should also be noted that some studies have reported positive effects of tasks on takeover performance. Naujoks et al. (2018) found visual and mental workload during engagement with a task reduced reaction times to a critical event. Likewise, Neubauer et al. (2012) found that phone use reduced reaction times, as measured via deacceleration to a critical event. Secondary tasks could therefore stop fatigue effects and prevent cognitive underload.

### 4.2.3 Mechanisms underlying performance deficits following non-driving related tasks

Many studies have utilised everyday tasks such as reading the news (Zeeb et al., 2017), interacting with an in-vehicle system (Toffetti et al., 2009), and watching a movie (Carsten et al., 2012). However, the discrepancy in results makes it unclear what aspects of a task negatively impact driving performance. Standardised tasks such as the working memory n-back task (Radlmayr et al., 2014) or an auditory oddball task (Körber et al., 2015) can provide insights into the mechanisms disrupted during engagement with secondary tasks. Radlmayr et al. (2014) found that the visual surrogate reference task increased collision rate, whereas an n-back task did not. Their results imply that visual tasks may recruit visual processing mechanisms that disrupt driving performance, whereas working memory mechanisms remain largely unaffected.

Gaze behaviours can provide additional insight into attentional allocation mechanisms impacted by secondary unrelated tasks. Ko & Ji (2018) found that time to watch the windshield and time to watch the road was greater following an n-back task when compared to a task subjectively rated as too easy or too difficult. However, Zeeb et al. (2017) found that time to watch the road did not differ between naturalistic tasks of writing an email, reading a news text, and watching a video clip. Although first gaze metrics can provide a measure of instant attentional allocation, it is important to understand how attention is modulated throughout driving as manual driving requires sustained attention. Merat et al. (2014) found that blink suppression was greatest (blink rate reduced) during a critical incident when following a takeover request from engagement with a prior task. Visual processing mechanisms may have been overcompensating to ensure successful takeover, and this effect lasted during the takeover process.

Physiological measures can provide insight into arousal-related mechanisms associated with takeover performance. Alrefaie et al. (2019) examined changes in heart rate during two different secondary tasks: emailing and twenty questions game. Participants displayed increased heart rate during the email task when compared to engaging in no task, and the twenty questions game when compared to no task. Takeover quality measures, PerSpeed (mean percentage change of vehicle speed for a 30 second period before the takeover) and PerAngle (mean percentage change of vehicle heading angle), were strongly correlated with heart rate, and linear mixed modelling revealed a 1% increase in heart rate corresponded to a 9.1% decrease in PerSpeed and a 54% change in PerAngle. Alrefaie et al.'s (2019) study provides the first support that heart rate is strongly associated with takeover performance, and suggests increased arousal is associated with worse driving performance.

### 4.2.4 Neurophysiological mechanisms during automated driving

Neurophysiological measures can detect effects of attention decrements during automated and manual driving. Solís-marcos et al. (2017) investigated attention-related event-related potentials while participants partook in an auditory oddball task during manual and semi-autonomous simulated driving. During semi-autonomous driving, P3b amplitude and subjective manual demand were lower, when compared to manual driving. The reduction in P3b amplitude represents a reduced ability to attend to target stimuli, providing support that attention was impaired during autonomous driving. EEG studies during manual driving have found that P3b amplitude reduces with increased fatigue (Schmidt et al., 2009; Zhao et al., 2012) and distractions (Karthaus et al., 2018). Analyses of oscillatory mechanisms of alpha have revealed parietal alpha event-related desynchronisation (ERD) during active driving (Sakihara et al., 2014; Schier, 2000), and is less pronounced during reduced vigilance (Correa et al., 2014; Huang et al., 2009). These studies provide some evidence for the recruitment of a posterior attention network during simulated driving, which could possibly be impacted during periods of automation. However, the mentioned studies focus on comparing EEG mechanisms during either only manual driving, or partially automated driving with no specific takeover. Therefore, they do not consider the impact of different secondary tasks on driving performance following a takeover request. It is unclear how oscillatory mechanisms related to attention are implicated during driving when following engagement of secondary tasks.

### 4.2.5 Experiment rationale

The discrepancy between results makes it unclear what aspects of a task negatively impact takeover performance and limited research has attempted to uncover the mechanisms related to driving performance following engagement with a task. Electrophysiological correlates of cognitive functioning have the potential to uncover the temporal evolution of such processes impacted. Therefore, the goal of the present study was to investigate how neural oscillatory mechanisms were altered during simulated driving when following a task differing in attentional load. This study was designed to understand how attention is modulated during takeovers in semi-autonomous driving. Similarly to Chapter 3, the present study is interested in whether the attentional demand of an unrelated task (i.e. how engaging a book is/perceptual load) has an impact on task (i.e. driving) performance. These sudden but vital task switches have life-threatening consequences and are made more challenging when humans have the potential to engage in non-related tasks.

A visual search task was administered as the unrelated task. Although several previous studies have administered naturalistic tasks, such as reading a book or writing emails, there are a number of limitations in doing so. Previous research has demonstrated a lack of consensus for the impact of type of task on takeover quality. In such studies, the authors assume that cognitive load is altered during different tasks. For example, in Dogan et al. (2019), the authors measured the effect of type of task by asking participants to switch between watching videos and writing emails. The authors assumed that watching videos would induce cognitive underload, whereas writing emails would increase load. The authors found no difference in takeover quality between the type of task. This may be because workload was not successfully manipulated. Therefore, using a standardised controlled task that has reliably been demonstrated to modulate certain mechanisms is beneficial to fully understand the impact on takeover quality. In addition, findings from Chapter 3 revealed that the visual search task modulated performance and oscillatory mechanisms successfully, further providing confidence that the visual search task can be used as a successful manipulation of attentional load. In addition to the visual search task, a passive viewing task, where participants were required to watch the roadway in front, was administered. This acted as a control condition as no physical task was undertaken, yet participants were still not in control of the vehicle.

The visual search task results should replicate the visual search task results in Chapter 3, demonstrating worse performance (i.e. increased reaction time, reduced accuracy), reduced stimulus processing (i.e. lower P3b amplitude), increased recruitment of attentional resources (i.e. frontal and parieto-occipital alpha desynchronisation (ERD)), and movement uncertainty (i.e. reduced frontal theta synchronisation (ERS)) during high load when compared to low load. Informed by the results from Chapter 3, and the previous research on manual driving (e.g. Correa et al., 2014; R. S. Huang et al., 2009; Sakihara et al., 2014; Schier, 2000), it was expected that engagement with search tasks would reduce performance and increase recruitment of cortical mechanisms required for processing of the simulated driving task. Behaviourally, driving performance indicated by longitudinal (speed) and lateral control, including time to acceleration, would suffer when following a search task compared to the passive viewing task, and more so following the high load search task when compared to the low load search task. Parietal-occipital alpha ERD coupled with an increase in evoked frontal theta ERS during simulated driving would be reduced significantly following the passive viewing task (control) when compared to both search tasks. These oscillatory mechanisms would be strongest following high perceptual load, representing a stronger recruitment of visuomotor attentional mechanisms.

## 4.3 Method

### 4.3.1 Participants

Thirty-eight healthy right-handed licensed drivers participated in this study (twenty-six females, twelve males, mean age ± SD = 23.22 ± 5.13 years, range 18-42). All participants considered themselves healthy as reported in self-reports. Due to the nature of the experiment, individuals with any neurological condition, hypersensitive skin, or skin allergies were excluded. As this study required visual and auditory processing, participants with uncorrected vision or hearing were excluded. Participants were all undergraduate students at the University of the West of England and were required to hold a driver's license so they had an understanding of general driving rules and vehicle control competencies. All participants gave written informed consent and were fully debriefed at the end of the study. Ethical approval was obtained by the Faculty of Health and Applied Sciences University of the West of England Research Ethics Committee (HAS.19.08.003).

Four participants were removed from all analyses due to technical errors (three) and an incomplete dataset (one). An additional participant was removed from all visual search analyses due to technical errors (EEG markers did not send across correctly), whereas a different participant was removed from all simulated driving analyses due to a technical error (EEG markers did not send across correctly). In the visual search analyses, thirty-three participants were included (twenty-two females, eleven males, mean age ± SD = 22.76 ± 6.01 years, range 18-42); in the simulated driving analyses, thirty-three participants were included (twenty-two females, eleven males, mean age ±SD = 23.11 ± 5.54 years, range 18-42).

### 4.3.2 Stimuli

Participants were seated in a STISIM Drive® (version 3) fixed-base driving simulator (Systems Technology Inc., Hawthorne, CA, USA). The driving environment was projected on to a 27-inch LCD monitor. To the right of the driving simulator, a Dell 17-inch LCD monitor with a 60 Hz refresh rate was presented at a viewing distance of 57 cm and was always closer to the participant than the driving simulator monitor (see Figure 4.1). This is similar to naturalistic settings where users may use technological devices, such as mobile phones and tablets, which are physically nearer to the user when compared to the driving environment. The Dellmonitor presented the visual search stimuli, considered here as the low load search task and high load search task. A keyboard was placed in front of the LCD monitor for participants to respond to visual search stimuli. Throughout the experiment, participants sat in the driving simulator seat. The simulator was set up comfortably so they were able to reach the pedals and steering wheel to be able to take control of the vehicle, as well as reach the keyboard and attend to the visual search stimuli on the 17-inch LCD monitor.

The experiment was carried out in six blocks of 30 trials. Each trial lasted approximately 19 seconds, while each block lasted ~10 minutes, with rest breaks in between. 60 trials in each condition (low load search task, high load search task, passive viewing task) were presented, resulting in 180 trials in total. A composite trial consisted of an auditory cue, the secondary task (either: passive viewing task, low load search task, high load search task), a second auditory cue, and simulated driving. An auditory cue presented at a comfortable listening level was presented to inform the participant which task to switch to: ‘Drive’, ‘Task’, or ‘Watch’. At the beginning of ‘Task’, participants were required to press the spacebar key to initiate the task. An initial fixation cross was presented on the screen for 1.5 s prior to the visual search presentations. The order of trials was randomised within each block; each non-driving related task was presented 10 times within a block and was not consecutively repeated more than twice over all blocks. See Figure 4.1 and Figure 4.2 for experimental set up and stimuli procedure.

#### *4.3.2.1 ‘Drive’*

The auditory cue ‘Drive’ instructed participants to take manual control of the driving simulator. STISIM3 software was used to present the visual environment and measure driving performance. The driving simulator included an adjustable car seat, gear stick, steering wheel, indicators, accelerator, clutch and brake pedals, dashboard, speedometer, as well as the right, left, and rear-view mirrors. The driving scenario consisted of a suburban three-lane environment, with cars on the road, and pedestrians on the pavement, with road signs and left carriageway driving consistent with the UK. Speed limit signs were included every 1000 ft in the environment. The STISIM driving simulator has been previously validated and a high transferability of simulator performance and on-road driving performance has been reported (e.g. Bédard et al., 2010; Lee, 2002). Participants were asked to adhere to the speed limit (60 mph) and keep to the left-hand lane with a steady lane position, as they normally would in real-world driving. The drive continued for 792 ft which lasted approximately nine seconds. Participants switched to ‘Drive’ after either ‘Task’ or ‘Watch’.

#### *4.3.2.2 ‘Task’*

The auditory cue ‘Task’ informed participants to switch to the visual search task. The stimuli used in the visual search task consisted of either low or high perceptual load and were the same stimuli used in Chapter 3 (see Chapter 3.3.2). In this experiment, visual search presentations in each trial were limited to seven. Participants did not have time to attend to all visual search presentations before the next auditory cue was presented. Therefore, participants were told to adhere to the cue instruction and move on. During ‘Task’, the driving simulator switched to automated driving and so the vehicle was in control of lateral position and acceleration. The vehicle kept to the lane position and kept speed constant at 60 mph.

*4.3.2.3 'Watch'*

The auditory cue 'Watch' required participants to watch the roadway in front. The vehicle switched to automated driving and so the vehicle was in control of lateral position and acceleration. The vehicle kept to the lane position and kept speed constant at 60 mph. This condition was considered the control condition, also described as the passive viewing task, as participants did not take part in a task before completing 'Drive'.

*Figure 4.1.* Experimental set up. Participants partook in simulated driving following a high load search task or a low load search task presented on a monitor, or a passive viewing task, where participants watched the roadway ahead.

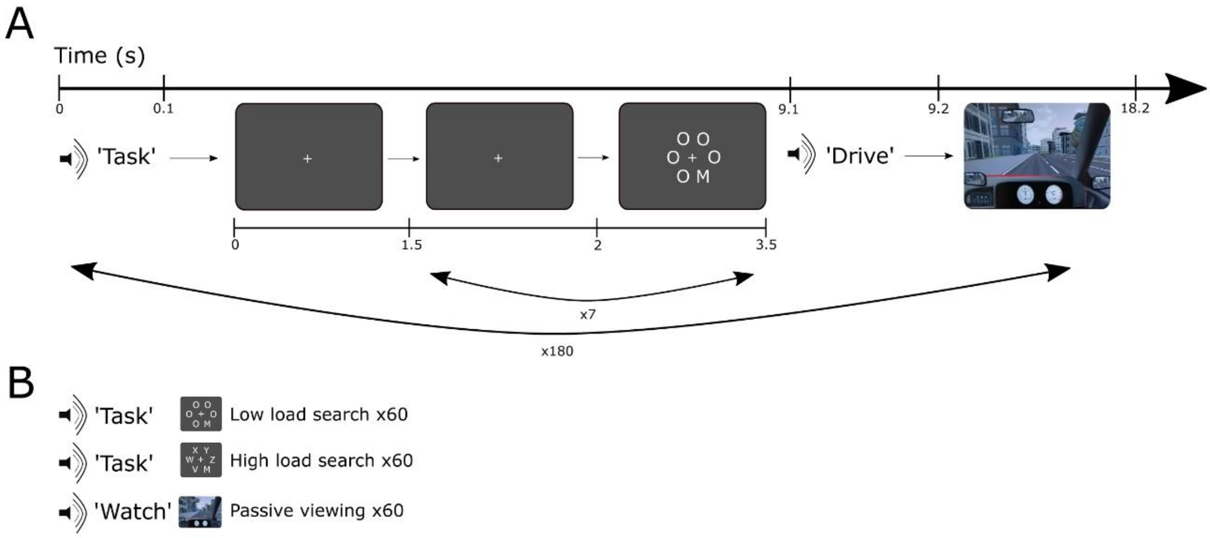



*Figure 4.2.* Schematic depiction of the experimental paradigm. **(A)** An example of an experimental trial. Participants switched between an unrelated task and manual driving. An auditory cue instructed the participant to either engage in a visual search task, 'Task', or a passive viewing task involving watching the roadway, 'Watch'. In this example, an auditory cue 'Task' was presented. During the task, visual search presentations were presented seven times. A further auditory cue instructed the participant to take control of the vehicle and participants drove for approximately nine seconds. 180 trials were presented in total. **(B)** Three conditions of the unrelated task. The task was either a visual search task of low perceptual load, a visual search task of high perceptual load, or a passive viewing task, where participants were instructed to watch the roadway. The two visual search tasks were preceded by an auditory cue 'Task'. The passive viewing task condition was preceded by an auditory cue 'Watch'. During this condition, the automated vehicle was in control for approximately nine seconds. Each condition was presented 60 times, randomised between trials.

### 4.3.3 EEG recordings

Continuous EEG was recorded using BrainVision actiCHamp 32-electrode EEG system with active electrodes. The ground electrode was located at Fpz (10-5 electrode system: Oostenveld & Praamstra, 2001). The online sampling rate was set at 1000 Hz and electrodes were referenced to Cz position. Impedances of all channels were adjusted to below 25 kOhms as recommended by the manufacturer. A StimTrak from BrainProducts was used to detect the onset of the auditory cue. An acoustical stimulator adapter was connected to the StimTrak and driving simulator PC sound card, to ensure that trigger generation marked the real onset of the auditory stimuli: 'Drive', 'Task', 'Watch'.

### 4.3.4 Protocol

Following on from prior screening and enrolment, participants were invited to attend the laboratory and on arrival, were introduced to the equipment used in the study and filled in a demographic questionnaire before providing consent. They were provided with verbal instructions for the task and completed a practice run consisting of six randomised trials, which always included two low load search tasks, two high load search tasks, and two passive viewing tasks conditions. The author observed their responses and made an informed decision whether they showed an objective understanding of the task. All participants showed an understanding of the task rules after the practice run and were happy to continue. Once the EEG set up was complete, all electrophysiological signals were checked visually before commencement of the experiment. A within-subject protocol required participants to switch between an unrelated task and simulated driving. The task was one of three: low perceptual load visual search, high perceptual load visual search, or a passive viewing task (control condition). The author stayed in the same room as the participant to ensure that they were carrying out the correct tasks. Any incorrectly responded to trials were recorded and later removed from subsequent analyses. Rest breaks occurred between blocks to allow participants to repose. Overall, participants were in the laboratory for around 2.15 hours: ~1hr set up and ~1.15 hour for the task.

### 4.3.5 Pre-processing

#### *4.3.5.1 Behavioural data*

Behavioural responses were collected for the visual search task and simulated driving. Pre-processing of behavioural data were undertaken using custom written scripts in MATLAB (MathWorks, Natick, MA). For the visual search task, accuracy and reaction time measures were calculated the same as Chapter 3 (see Chapter 3.3.5.1). As some visual search presentations were not responded to as an auditory cue had already been presented, any non-response after five sequential presentations were not included in the accuracy analyses. Takeover time was derived from the time the auditory takeover was announced, until the time it took the participant to engage with the accelerator pedal. This was because the speed of the vehicle rapidly decreased once the takeover was initiated, and so participants were required to engage with the accelerator pedal as quickly as possible during simulated driving. Longitudinal acceleration with a corresponding time stamp was exported from STISIM3. The time the participant engaged with the accelerator pedal was subtracted from the time the auditory cue was presented, for each trial. Takeover quality was captured by measures of standard deviation of the lateral position (SDLP), standard deviation of vehicle speed (SDS), and vehicle speed, to index lateral and longitudinal control, respectively. Lateral position and vehicle speed were exported directly from STISIM3. Only correctly responded to trials were used. Driving performance was averaged across trials.

#### *4.3.5.2 EEG data*

Pre-processing of EEG data was performed using EEGLAB (Delorme & Makeig, 2004) and custom-written scripts in MATLAB (The MathWorks, Natick, MA, USA). To begin, the auditory markers were edited to represent the task (low load search, high load search, passive viewing) to allow for subsequent epoching. The data were pre-processed similarly to Chapter 3 (see Chapter 3.3.5.2). Firstly, the data were downsampled to 500 Hz due to RAM requirements. Next, the data were bandpass filtered using an FIR filter between 0.1 to 100 Hz.

Data were then epoched into three separate sets: visual search, simulated driving (including the passive viewing condition), and baseline for simulated driving. For the visual search task, the data were epoched from -1.5 to 2 s in relation to visual search onset (0s). The data were linearly baseline corrected from -200 to 0 ms. Only correctly responded to trials were analysed. The time periods analysed were the same as Chapter 3 and was explored to ensure that participants demonstrated the expected EEG differences during low and high visual search load. For simulated driving, the 9 s driving period was of interest. Therefore, the data were epoched from -3 to 13 s in relation to the auditory cue onset 'Drive' (0 s). A final dataset was created for the baseline in relation to the simulated driving. Data were epoched from -2.5 to 3.5 s in relation to the initial fixation cross at the start of each trial. As participants had to initiate the fixation cross presentation, this data was importantly not contaminated by significant movement or task activity. Therefore, this time period provided an appropriate baseline. Any incorrectly answered responded to trials were removed at this point. For all three epoched datasets, manual trial rejection was undertaken to extract major sources of artefacts. Independent components analysis was computed for further artefact rejection. Once artefacts were removed, the data were re-referenced to the average of all EEG channels.

*4.3.5.2.1 Event-related potentials (ERPs)*

To capture visual search event-related potentials (ERPs), the data were further filtered between 0.3 to 35 Hz. The data were then epoched from -200 to 1000 ms, and condition-averaged to create subject specific ERPs. To investigate the impact of perceptual load on attentional allocation, mean amplitude of the P3b was extracted for each visual search presentation during low and high perceptual load. As informed by Chapter 3, a latency window of 300 – 450 ms was used over electrode Pz (see Chapter 3.3.5.2.1). To investigate the impact of perceptual load on perceptual processing, mean amplitude of the N1 component for each participant was taken in a 50 ms latency window around the peak of the N1 (peak = 170 ms; range = 145 – 195 ms). Data from electrodes O1, O2 and OZ were extracted and averaged together.

*4.3.5.2.2 Spectral analyses*

To capture spectral oscillatory data, the re-referenced visual search data were epoched into low and high load search conditions. The epoched simulated driving data was epoched regarding which task preceeded it: low load search task, high load search task, and passive viewing task. Another epoch was created for 'passive viewing' and consisted of the data during 'Watch', where participants observed the roadway in front. Additional exploratory analyses were carried out on this epoch to compare attention-related alpha activity during passive viewing and manual driving (see results Chapter 4.4.2).

All epoched data was decomposed into a time-frequency representation with linear scaling for 150 frequencies between 3 and 80 Hz from fast Fourier transform and via Morlet complex wavelet convolution, followed by the inverse fast Fourier transform. Cycles were increased with frequency, starting at 3 and going up in steps of 0.5. The wavelet transform was performed for each trial. Power ($\mu V^2$) of oscillatory activity was computed. In order to remove scale differences between individuals, all power values in the time-frequency representation were normalised by decibels to the baseline power (dB). For the visual search task, baseline power was computed as the average power from -400 to -200 ms pre-stimulus at each frequency band. For simulated driving, the baseline power was computed as the average power from 500 to 1500 ms post-stimulus in the baseline epoch. A condition-average baseline was used for several reasons. Firstly, the signal-to-noise ratio of the baseline is increased (compared to a condition-specific baseline) as both low and high load fixation cross periods were included in the baseline (Cohen, 2014). This was important as only 60 trials in each condition were administered due to extensive trial length resulting in prolonged data collection sessions (~19 s rather than 2 to 3 s in conventional EEG studies). Secondly, the passive viewing condition did not have an appropriate baseline period, as participants focused on the roadway as soon the auditory cue was presented. Therefore, the same baseline was applied to each condition's time-frequency representation. Ultimately, the absolute values of the resulting transforms were trial averaged to provide a time-frequency decomposition for each participant over every electrode.

Time-frequency regions were defined based on a priori hypotheses from the previous experiment described in Chapter 3. As no lateralisation effects were expected during simulated driving, parieto-occipital, rather than parietal only, attention-related oscillatory data were averaged together. Visualisation of the condition-averaged results provided an exact time window and exact frequency range to extract data from (Cohen, 2014). The activity from all pixels in a region of interest was averaged for each participant. See Appendix 4.1 and Appendix 4.2 for condition grand-averaged time-frequency spectrograms.

*Visual search:* Regions of interest were defined for alpha (10.5 - 12.5 Hz) and theta (4 – 7 Hz) power. Frontal activity was averaged over electrodes Fz, FC1, FC2 between time 200 – 600 ms for alpha, and 200 - 600 ms for theta. Parieto-occipital activity was averaged over electrodes O1, O2, P3, P4, P7, P8 between time 200 – 600 ms for alpha and 100 – 300 ms for theta.

*Simulated driving:* Regions of interest were defined for theta (4 – 7 Hz), alpha (10.5 – 12.5 Hz) and beta (15 – 25 Hz). Frontal theta activity was averaged over electrodes Fz, FC1, FC2 between time 200 – 900 ms. Frontal alpha activity was averaged over electrodes Fz, FC1, FC2 between 900 – 9100 ms. Parieto-occipital alpha activity was averaged over electrodes P3, P7, P4, P8, O1, O2 between time 900 – 9100 ms. Parieto-occipital theta activity was averaged over electrodes P3, P7, P4, P8, O1, O2 between time 200 – 800 ms. Motor-related activity was extracted and averaged over electrodes C3 and C4 between time 950 – 9100 ms for beta and alpha power.

### 4.3.7 Statistical analyses

All statistical analyses were performed using IBM SPSS Statistics for Windows, version 26 (IBM Corp., Armonk, N.Y., USA). Assumptions of sphericity were tested using Mauchly's test, and if violated, Greenhouse-Geisser estimates were used in the repeated measures calculations. The statistical threshold for significance was set to two-tailed $p < .05$. Effect size was reported as eta squared ($\eta^2$) for one-way ANOVA significant results, and Cohen's $d_z$ was reported for paired-samples $t$-tests.

*Behavioural analyses.* Paired-samples *t*-tests were used to compare accuracy and response time measures between visual search low and high load conditions. For time to accelerator pedal and vehicle speed, a one-way repeated measures ANOVA with repeated measures (Task: passive viewing, low load search, high load search) was undertaken. Shapiro-Wilk test of normality and visualisation of QQ plots of the unstandardized residuals indicated that SDS and SDLP were not normally distributed. SDS data were normalised via the natural logarithm before the one-way repeated measures ANOVA was undertaken. For SDLP, the non-parametric Friedman test was undertaken on task (Task: passive viewing, low load search, high load search). This was due to the severe skewedness of the data which could not be resolved by logarithmic transformations. *Post hoc* analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < .017$ (i.e. minimum significance level / number of tests = .05 / 3). Effect size *r* was calculated by dividing the test statistic by the square root of the number of observations (Field, 2013, p. 257; Pallant, 2007, p. 225).

*EEG analyses.* Paired-samples *t*-tests were used to compare alpha and theta oscillations between visual search low and high load conditions. For simulated driving spectral analyses, a one-way repeated measures ANOVA was performed (Task: passive viewing, low load search, high load search). Exploratory analyses on oscillatory activity between passive viewing and automated driving compared to manual driving was conducted with a paired-samples *t*-test. Further analyses were undertaken with a series of paired-samples *t*-tests on every time point (length between time points = 74 ms). Multiple comparisons were controlled for via the Benjamini & Hochberg (1995) procedure.

*Correlation analyses.* Participant by participant correlations were performed on significant continuous driving performance measures (takeover quality) and parieto-occipital alpha activity. Six repetitions (relating to two significant driving performance measures in three conditions: passive viewing, low load search, high load search) of the correlation analysis gave a Bonferroni corrected statistical threshold of $p < .008$. A Pearson's correlation coefficient was calculated for vehicle speed, whereas the non-parametric Spearman Rho was calculated for SDLP.

## 4.4 Results

### 4.4.1 Behavioural performance

For visual search, accuracy and reaction time values were analysed and compared with paired-samples *t*-tests between low load and high load. Simulated driving over 9 seconds produced four measures in total. One instantaneous measure was derived to infer takeover time: time to accelerator pedal, and three continuous measures were derived to infer takeover quality: standard deviation of lateral position (SDLP), standard deviation of vehicle speed (SDS), and overall vehicle speed. Each measure was calculated and analysed with a one-way ANOVA with the within-subject factor Task: simulated driving following passive viewing, simulated driving following visual search low load, and simulated driving following visual search high load.

#### *4.4.1.1 Visual search task performance*

A paired-samples *t*-test revealed a significant difference in accuracy between the low load and high load conditions, $t_{(33)} = 8.36$, $p < .001$, $d_z = 1.43$. As expected, accuracy was greater during low load. Similarly, there was a significant difference in reaction time between low load and high load, $t_{(33)} = -16.36$, $p < .001$, $d_z = 2.81$. As expected, reaction time was greater during visual search high load. Overall, reaction time and accuracy measures reveal worse performance during visual search high load when compared to low load. See Figure 4.3 and Table 4.1 for means and SDs.

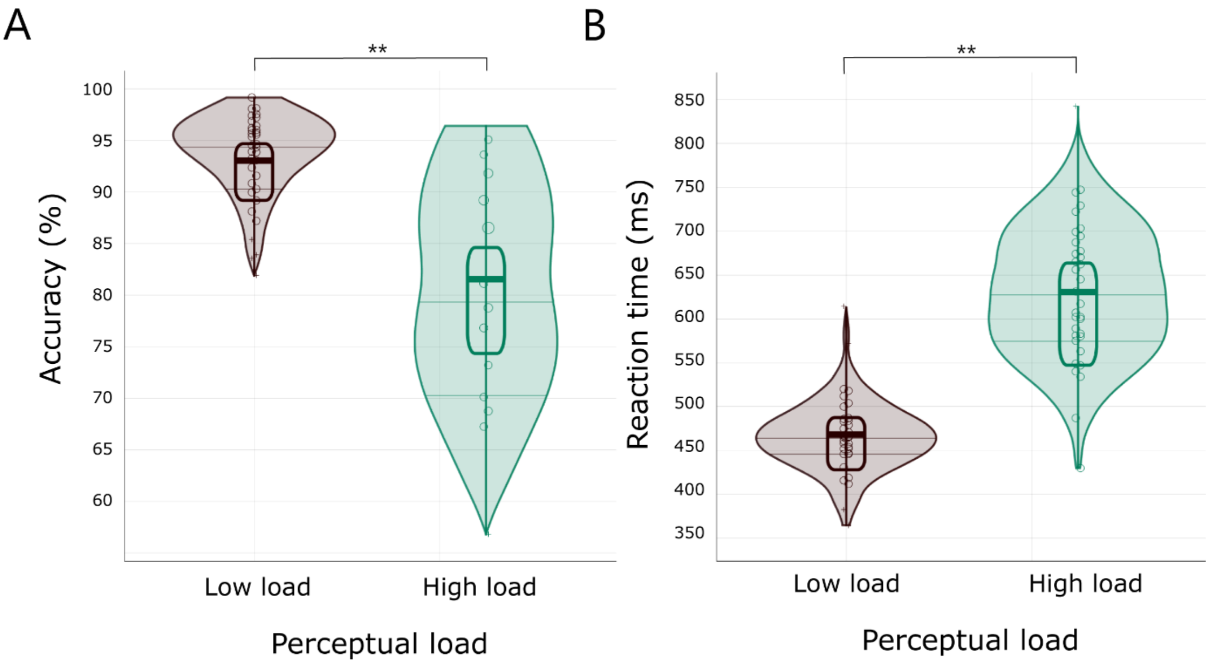



*Figure 4.3.* Violin plots for visual search task performance data during low and high perceptual load. **(A)** Mean accuracy score (%) **(B)** Mean reaction time (ms). Violin plots represent the distribution of each data series, with a box plot and whisker drawn over these data. Key: ** $p < .001$; notch centre = mean

### *4.4.1.2 Simulated driving performance*

For takeover time, a one-way repeated measures ANOVA revealed that time to accelerator pedal was statistically significant between Tasks, $F_{(2,\ 62)} = 18.34, p < .001, \eta^2 = 0.38$. Pairwise comparisons revealed that time to accelerator pedal was greater when driving was preceded by a search task of low load compared to passive viewing task ($p < .001$), and when preceded by a task of high load compared to passive viewing task ($p < .001$).

For takeover quality metrics, a one-way repeated measures ANOVA revealed that vehicle speed was statistically significant between Tasks, $F_{(1.57,\ 51.83)} = 11.74, p < .001, \eta^2 = 0.26$. Pairwise comparisons revealed that vehicle speed deviated from the limit (60 mph) more when driving was preceded by a task of low load compared to passive viewing task ($p < .001$), and when preceded by a task of high load compared to passive viewing task ($p < .003$). Speed was lower following both low load and high load visual search when compared to passive viewing task. A one-way repeated measures ANOVA revealed that standard deviation of speed was statistically significant between Tasks, $F_{(1.54,\ 49.3)} = 4.16$, $p$

$= 0.03$, $\eta^2 = 0.12$. However, pairwise comparisons did not survive Bonferroni corrections. A statistically significant difference was also found between Tasks for standard deviation of lateral position, $\chi^2_{(2)} = 41.46$, $p < .001$. Wilcoxon signed rank tests revealed there were no significant differences between low load search and high load search, $Z = 1.76$, $p = .08$. However, there was a statistically significant increase in SDLP following low load search task compared to passive viewing task, $Z = -4.66$, $p < .001$, r = - .81, and following a high load search compared to passive viewing, $Z = -4.19$, $p < .001$, r = -.72. See Figure 4.4 and Table 4.1 for an overview of results.

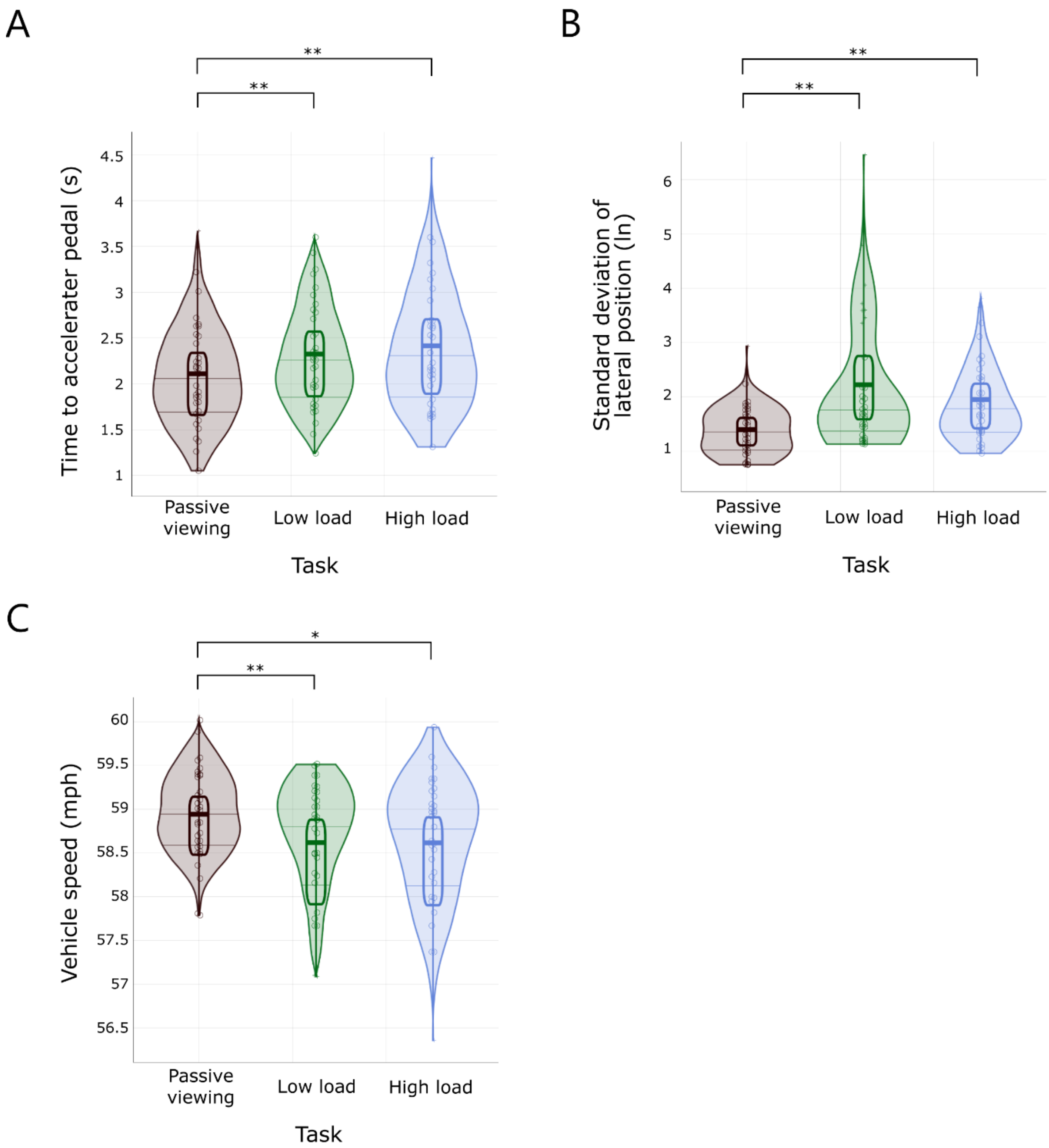



*Figure 4.4.* Violin plots of driving performance following engagement in tasks: passive viewing, low load visual search, high load visual search. **(A)** Average time to accelerator pedal (s). **(B)** Average standard deviation of lateral position (ln) **(C)** Average vehicle speed (mph). Key: * $p < .05$; ** $p < .001$; notch centre = mean.

*Table 4.1.* Visual search and simulated driving task performance. Mean (SD) are presented for all excluding standard deviation of the lateral position (SDLP), where median (IQR) is presented.

| Task | Measure | Task | | |
|---|---|---|---|---|
| | | Passive viewing | Low load search | High load search |
| **Visual search** | Accuracy (%) | - | 93.04 (4.56) | 81.54 (8.55) |
| | Reaction time (ms) | - | 468 (47) | 630 (83) |
| **Simulated driving** | Time to accelerator pedal (s) | 2.06 (0.52) | 2.28 (0.56) | 2.38 (0.68) |
| | Vehicle speed (mph) | 58.94 (0.52) | 58.62 (0.71) | 58.62 (0.52) |
| | Standard deviation of lateral position (SDLP; ft) | 1.32 (0.74) | 1.71 (2.04) | 1.84 (1.21) |
| | Standard deviation of vehicle speed (SDS; ln) | 3.37 (0.01) | 3.37 (0.02) | 3.38 (0.01) |

### 4.4.2 Passive viewing versus manual driving

Next, EEG data was analysed to explore the fluctuations in attention-related parieto-occipital alpha during the 9 s simulated driving period, when compared to passive viewing where participants monitored the roadway ahead. Two additional participants were removed from the analyses due to technical errors (‘Watch’ event marker failed to send across) and therefore 31 participants were run in total.

A paired-samples *t*-test revealed that parieto-occipital alpha desynchronisation (ERD) did statistically differ between passive viewing and manual driving, $t_{(31)} = -2.52$, $p = .02$, $d_z = -0.45$. Alpha ERD was greater during manual driving ($M = -2.84$, $SD = 2.92$), compared to passive viewing ($M = -1.87$, $SD = 2.19$). Additional analyses of paired-samples *t*-tests on each time point revealed alpha desynchronisation was greater at the beginning of the trial from the onset at alpha 998 until 4138 ms, and significant at time point 4512 ms, and between 5036 to 5186 ms. These results demonstrate greater alpha ERD during driving task engagement, compared to viewing the driving environment only. See Appendix 4.3 for *t*-test results and Figure 4.5 for visualisation of results.

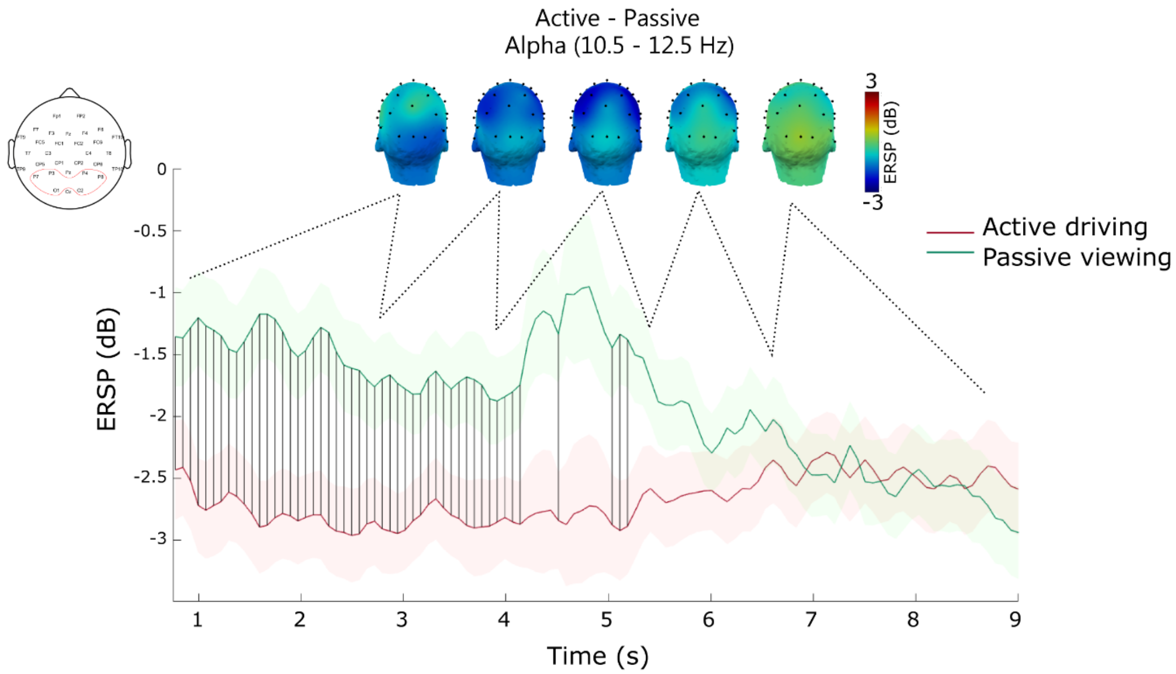



*Figure 4.5.* Grand-averaged dB normalised signal of parietal alpha (10.5 – 12.5 Hz) during passive viewing of the driving environment and manual driving. Significant time points are depicted by a grey line. Shaded areas represent the ± standard error of the mean difference. Topographical difference (active – passive) plots represent alpha activity averaged over five time series: 900 – 2540 ms, 2540 – 4180 ms, 4180 – 5820 ms, 5820 – 7460 ms, 7460 – 9100 ms.

### 4.4.3 Visual search EEG

Next, EEG data from the visual search task was analysed. Analyses were replicated from Chapter 3, to ensure that the visual search task modulated attentional load. First, event-related potential (ERP) components the N1 and P3b were compared with paired-samples *t*-tests between low and high visual search load. Next, spectral analyses were undertaken to explore alpha and theta dynamics between low and high visual search load. A series of *t*-tests were undertaken to explore frontal alpha power, frontal theta power, parieto-occipital alpha power, and parieto-occipital theta power, between low and high load visual search tasks.

#### *4.4.3.1 Event-related potentials (ERPs)*

A paired-samples *t*-test revealed that N1 amplitude did not statistically significantly differ between low and high load, $t_{(33)} = -.45$, $p = .84$. N1 amplitude was similar during low load ($M = -2.17$, $SD = 2.90$)

and high load ($M$ = -2.21, $SD$ = 3.19). A paired-samples $t$-test revealed P3b amplitude significantly differed between low and high perceptual load, $t_{(33)} = 3.72$, $p < .001$, $d_z = 0.64$. P3b amplitude was greater during low load ($M$ = 3.84, $SD$ = 4.72) when compared to high load ($M$ = 1.75, $SD$ = 3.19). See Figure 4.6.

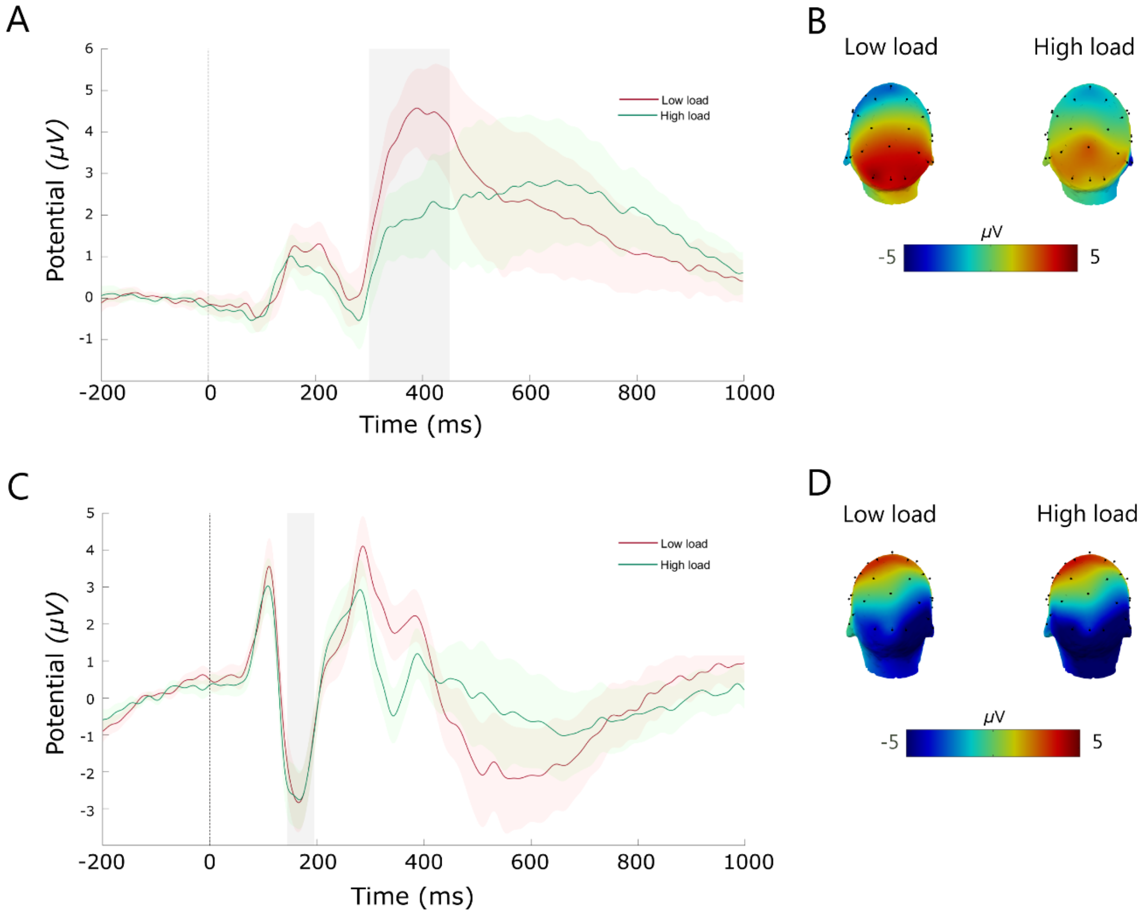



*Figure 4.6.* Grand-averaged P3b and N1 ERPs time-locked to visual search presentation under low and high perceptual load. Shaded areas represent the time course extracted for statistical analyses. Shaded error bars represent the ± standard error of the mean difference. **(A)** Grand-averaged P300. Correct trials were averaged across midline parietal electrode Pz. **(B)** P3b (300 – 450ms) mean amplitude scalp topography for low and high perceptual load. **(C)** Grand-averaged N1. Correct trials were averaged across occipital sites: O1, O2, OZ. **(D)** N1 (145 – 195ms) mean amplitude scalp topography for low and high perceptual load.

*4.4.3.2 Spectral analyses*

A paired-samples *t*-test revealed frontal alpha desynchronisation was significantly greater during high compared to low load, $t_{(33)} = 2.83$, $p = .008$, $d_z = 0.49$. Similarly, parieto-occipital alpha desynchronisation was significantly greater during high compared to low load, $t_{(33)} = 3.67$, $p < .001$, $d_z = 0.63$. A paired-samples *t*-test revealed frontal theta synchronisation was greater during low load compared to high load, $t_{(33)} = 2.57$, $p = .02$, $d_z = 0.44$. However, parieto-occipital theta synchronisation did not differ between low or high load, $t_{(33)} = -0.73$, $p = .47$. See Table 4.2 for an overview of the means and standard deviations and Figure 4.7 for a graphical representation of results.

Altogether, the visual search EEG results replicate Chapter 3. High perceptual load reduced P3b amplitude while reducing frontal theta ERS, increasing frontal alpha ERD, and increasing parieto-occipital alpha ERD. See Table 4.3 for an overview of behavioural and EEG visual search results.

*Table 4.2.* Mean (SD) of visual search time-frequency power (dB normalised).

| Frequency (Hz) | Averaged electrodes | Perceptual load | |
|---|---|---|---|
| | | Low load | High load |
| Alpha (10.5 – 12.5)* | Fz, FC1, FC2 | -1.27 (1.39) | -1.6 (1.32) |
| Alpha (10.5 – 12.5) ** | P3, P7, P4, P8, O1, O2 | -2.31 (1.47) | -2.81 (1.47) |
| Theta (4 – 7)* | Fz, FC1, FC2 | 1.21 (0.72) | 0.90 (0.84) |
| Theta (4 – 7) | P3, P7, P4, P8, O1, O2 | 1.30 (1.13) | 1.42 (1.43) |

*Key:* * $p < .05$; ** $p < .001$

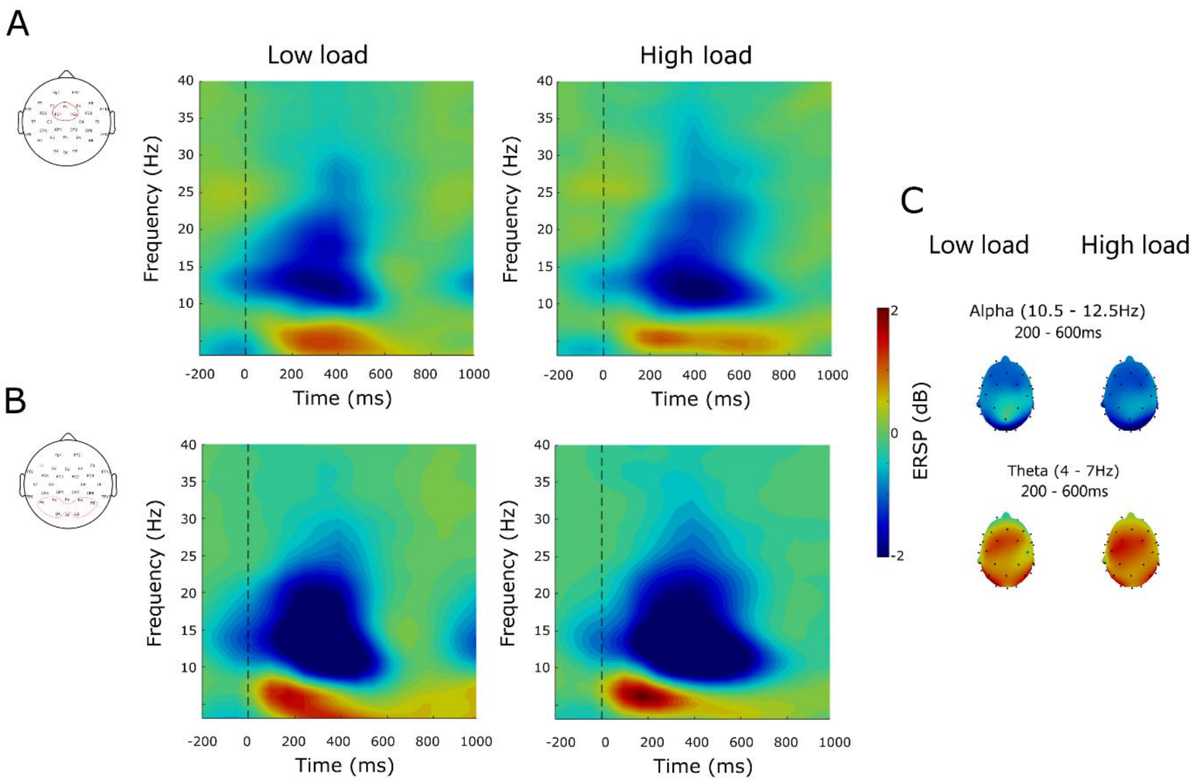



*Figure 4.7.* Time-frequency results for visual search activity over low and high perceptual load. **(A)** Grand-averaged time-frequency spectrograms for frontal (FC1, FC2, Fz) visual search activity. **(B)** Grand-averaged time-frequency spectrograms for parieto-occipital (P4, P7, P8, P3, O1, O2) visual search activity. **(C)** Averaged topographical plots over alpha 200 – 600 ms and theta 200 – 600 ms.

*Table 4.3.* Overview of visual search results.

| Measure | Parameter | Time window (ms) | High compared to low load |
|---|---|---|---|
| **Behavioural** | Accuracy (%) | – | ↓ |
| | Reaction time (ms) | – | ↑ |
| **EEG** | N1 (μV) | 145 – 195 | – |
| | P3b (μV) | 300 – 450 | ↓ |
| | Frontal alpha ERD (dB) | 200 – 600 | ↑ |
| | Parietal-occipital alpha ERD (dB) | 200 – 600 | ↑ |
| | Parietal-occipital theta ERS (dB) | 100 – 300 | – |
| | Frontal theta ERS (dB) | 200 – 600 | ↓ |

*Key:* '↑' represents the parameter was significantly greater during high load, '↓' was significantly lower during high load, and '–' represents the parameter was similar between low and high load. Key: event-related synchronisation: ERS; event-related desynchronisation: ERD; dB: decibel normalised

### 4.4.4 Simulated driving EEG

Finally, EEG activity was analysed over the 9 second simulated driving task. Analyses were interested in the impact of the attentional load of the preceding visual search task, and so the data compared were simulated driving following passive viewing, simulated driving following low load search, and simulated driving following high load search. Frontal theta, parieto-occipital theta, frontal alpha, and parieto-occipital alpha were analysed. Similar to Chapter 5's tracking task, theta activity was a transient response apparent in the first second of the task and was therefore associated with the task switch. Alpha was apparent over the task period, lasting approximately 9 seconds, representing the temporal dynamics of attentional demand. One-way (Task: passive viewing, low load search, high load search) repeated measures ANOVAs were undertaken on oscillatory activity.

#### *4.4.4.1 Theta activity*

A one-way repeated measures ANOVA revealed that frontal theta synchronisation was statistically significant between Tasks, $F_{(2, 66)} = 11.38$, $p < .001$, $\eta^2 = 0.23$. Theta synchronisation was lowest after the passive viewing task, compared to the low load search task ($p = .003$) and a high load search task ($p < .001$). Similarly, a one-way repeated measures ANOVA revealed that parieto-occipital theta synchronisation was statistically significant between Tasks, $F_{(1.43, 47.25)} = 13.32$, $p < .001$, $\eta^2 = 0.29$. Theta synchronisation was lowest after a passive viewing task, compared to a low load search task ($p = .003$) and a high load search task ($p < .001$). See Figure 4.8 and Table 4.4.

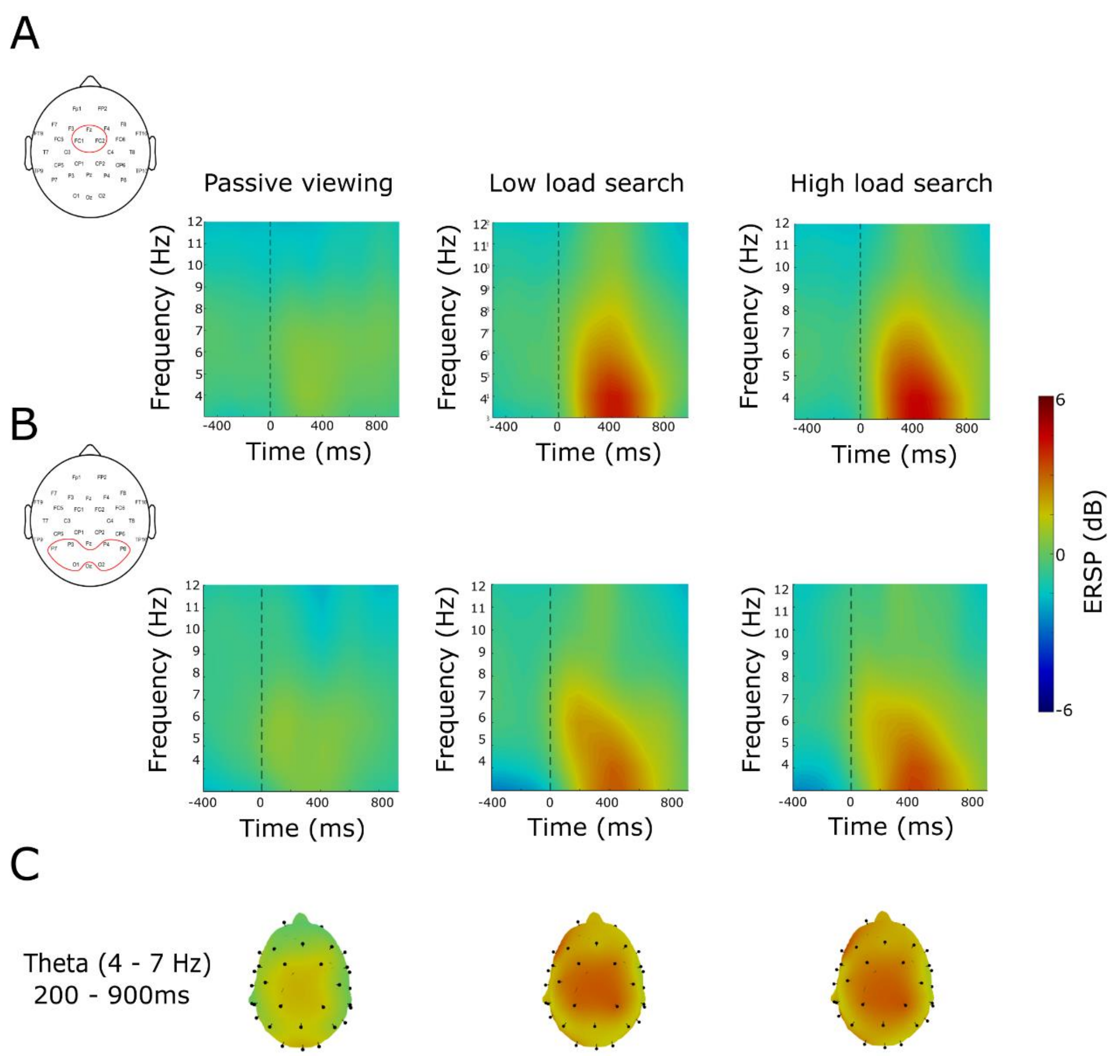



*Figure 4.8.* Theta (4 - 7Hz) activity during simulated driving following a passive viewing task, low load search task, and high load search task. **(A)** Grand-averaged time-frequency spectrogram of frontal theta activity. **(B)** Grand-averaged time-frequency spectrogram of parieto-occipital theta activity. **(C)** Topographical plots representing theta averaged between 200 – 900 ms post cue onset.

*4.4.4.2 Alpha activity*

A one-way repeated measures ANOVA revealed that frontal alpha desynchronisation was not statistically significant between Tasks, $F_{(2,\ 66)}$ = 1.14, $p$ = .33. A further one-way repeated measures ANOVA was performed on parieto-occipital activity during simulated driving. The ANOVA revealed a main effect of Task, $F_{(2,\ 66)}$ = 4.58, $p$ = .01, $\eta^2$ = 0.12. *Post hoc* comparisons revealed that a high load search task significantly differed from a passive viewing task. Alpha desynchronisation was greater following passive viewing search task compared to the high load search task ($p$ = .03). See Figure 4.9 for a graphical representation and Table 4.4 for means and standard deviations.

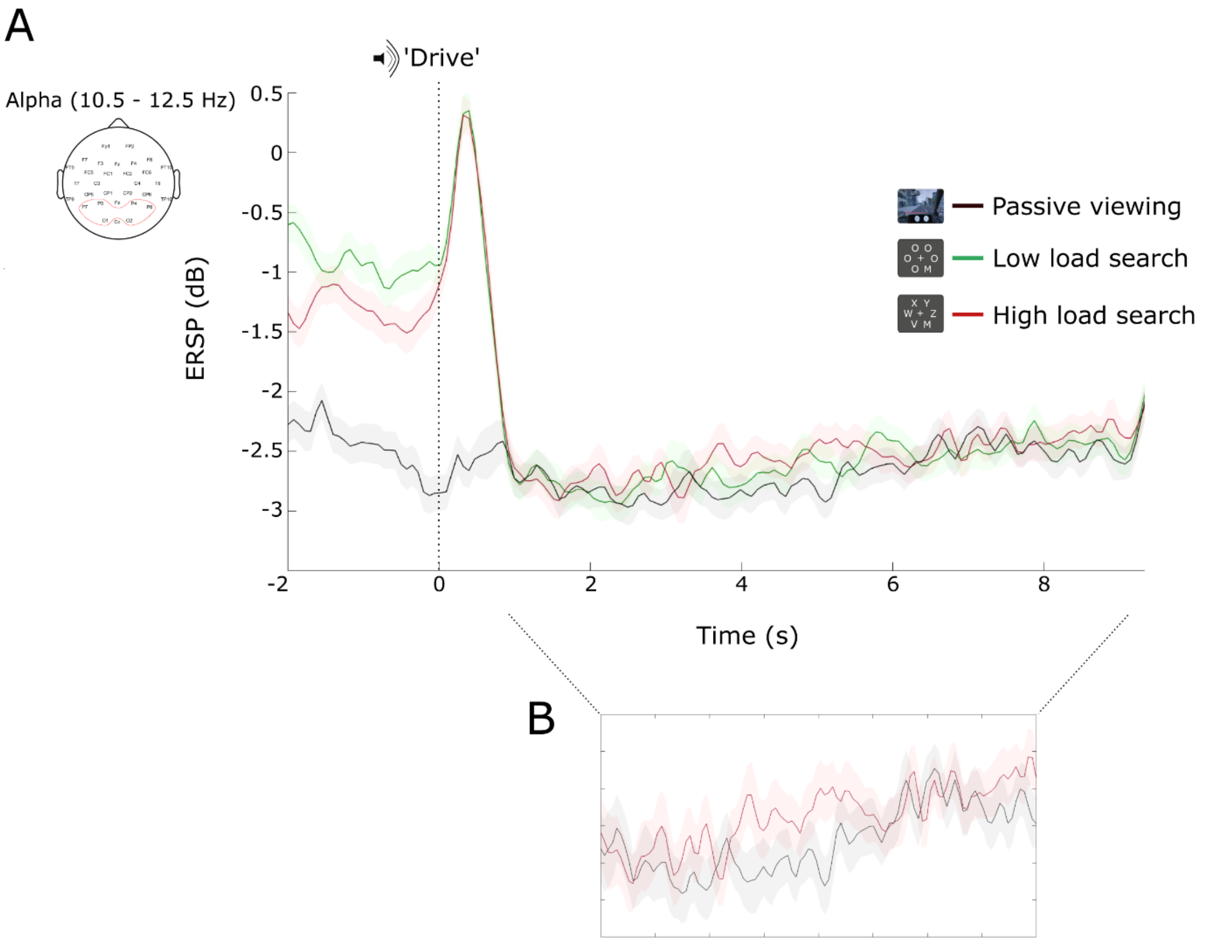



*Figure 4.9.* dB normalised signal for parieto-occipital alpha (10.5 – 12.5 Hz) during simulated driving. **(A)** Grand-averaged dB normalised signal for parieto-occipital alpha (10.5 – 12.5 Hz) during simulated driving following low load search task, high load search task, and passive viewing task. Time point 0 s depicts the auditory cue 'Drive'. Shaded error bars represent the ± standard error of the mean difference. **(B)** Close-up of passive viewing task and high load search task during simulated driving. Alpha ERD was significantly greater following the passive viewing task.

*4.4.4.3 Motor activity*

Control analyses were run to ensure that motor activity (C3 and C4) did not differ between tasks. A one-way repeated measures ANOVA on beta activity found no significant differences between Tasks, $F_{(1.34,\ 42.72)}$ = 1.49, $p$ = .23. A further one-way repeated measures ANOVA revealed there was no significant effect of motor-related alpha on Task, $F_{(1.42,\ 45.54)}$ = .29, $p$ = .76. Table 4.4 provides means and standard deviations for motor-related oscillatory activity.

*Table 4.4.* Mean (SD) of simulated driving time-frequency power (dB normalised).

| Frequency (Hz) | Averaged electrodes | Task | | |
|---|---|---|---|---|
| | | Passive viewing | Low load search | High load search |
| Theta (4 – 7) | FC1, FC2, Fz | 0.76 (3.22) | 1.83 (3.89) | 2.11 (3.75) |
| Theta (4 – 7) | P4, P8, P3, P7, O1, O2 | 1.04 (3.39) | 2.55 (3.65) | 2.78 (3.80) |
| Alpha (10.5 – 12.5) | FC1, FC2, Fz | -1.73 (2.81) | -1.68 (2.90) | -1.65 (2.86) |
| Alpha (10.5 – 12.5) | P4, P8, P3, P7, O1, O2 | -2.67 (2.92) | -2.60 (2.96) | -2.54 (2.91) |
| Beta (15 – 25) | C3, C4 | -1.43 (1.80) | -1.23 (2.07) | -1.23 (2.18) |
| Alpha (10.5 – 12.5) | C3, C4 | -1.36 (2.40) | -1.31 (2.50) | -1.31 (2.50) |

### *4.4.3.4 Correlations between parieto-occipital alpha and driving performance*

In regards to takeover quality metrics, representing continuous measures of performance, significant differences between vehicle speed and the standard deviation of lateral position (SDLP) were found between conditions, and therefore a participant by participant correlation was undertaken on vehicle speed with parieto-occipital alpha activity, and SDLP and parieto-occipital alpha activity, for each condition ($p$ < .008 corrected). A Pearson's correlation revealed that parieto-occipital alpha negatively correlated with vehicle speed following the low load search task, $r$ = -.46, $p$ = .008, and following high load search task, $r$ = .51, $p$ = .002. The correlation between alpha and a passive viewing task did not reach statistical significance, $r$ = -.3 $p$ = .09. Overall, following a low load or high load search task, greater alpha ERD represented better performance (closer to the speed limit: 60 mph). See Figure 4.10. The non-parametric Spearman's rank correlation revealed no significant association between SDLP and alpha during simulated driving following a low load search task, $r_s$ = -.01, $p$ = .95, a high load search task, $r_s$ = .01, $p$ = .95, or a passive viewing task, $r_s$ = .04, $p$ = .83.

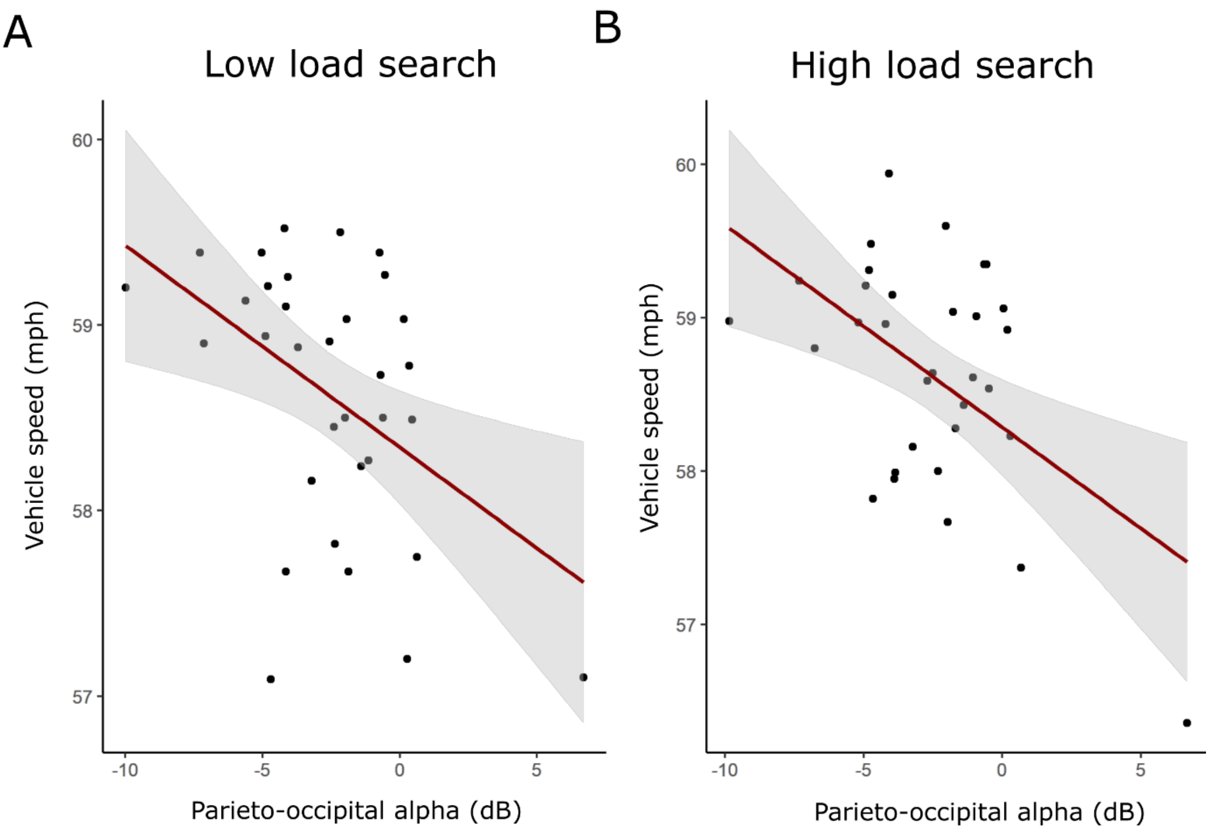



*Figure 4.10.* Significant correlations between parieto-occipital alpha and vehicle speed. Linear trendline depicted as a red line with shaded area representing 95% confidence interval. **(A)** Parieto-occipital alpha negatively correlated with vehicle speed following a low load search task. **(B)** Parieto-occipital alpha negatively correlated with vehicle speed following a high load search task.

*4.4.3.5 Overview of simulated driving results*

Table 4.5 provides an overview of all simulated driving results. Performance measures revealed that both takeover time and takeover quality were impacted by the engagement of a previous unrelated search task. Time to accelerator pedal was greater (i.e. increased reaction time) following a low visual search task when compared to a passive viewing task, and following a high visual search task when compared to a passive viewing task. Longitudinal (speed) control was also negatively impacted, and standard deviation of lateral position was greater, following both the visual search low and high task when compared to the passive viewing task. Similarly, transient frontal ERS and parieto-occipital theta ERS was greater following both low and high search task when compared to the passive viewing task. Parieto-occipital alpha ERD was lower following a high load search task when compared to a passive viewing task.

*Table 4.5*. Overview of simulated driving results.

| Measure | Parameter | Time window (ms) | Low – High | Passive viewing – Low | Passive viewing – High |
|---|---|---|---|---|---|
| **Behavioural** | Time to accelerator pedal (s) | | – | Low ↑ | High ↑ |
| | Vehicle speed (mph) | | – | Low ↓ | High ↓ |
| | Standard deviation of the lateral position (ln) | | – | Low ↑ | High ↑ |
| | Standard deviation of vehicle speed (mph) | | – | – | – |
| **EEG** | Frontal theta ERS (dB) | 200 – 900 | – | Low ↑ | High ↑ |
| | Parieto-occipital theta ERS (dB) | 200 – 800 | – | Low ↑ | High ↑ |
| | Frontal alpha ERD (dB) | 900 – 9100 | – | – | – |
| | Parieto-occipital alpha ERD (dB) | 900 – 9100 | – | – | High ↓ |
| | Motor beta ERD (dB) | 950 – 5100 | – | – | – |
| | Motor alpha ERD (dB) | 950 – 5100 | – | – | – |

*Key:* '↑' represents the parameter was statistically significantly greater during the specified condition, '↓' was significantly lower during the specified condition, and '–' represents the parameter was similar between both conditions. Key: event-related synchronisation: ERS; event-related desynchronisation: ERD; dB: decibel normalised

## 4.5 Discussion

The aim of this study was to investigate the modulation of task-related and attentional-related oscillatory mechanisms, namely theta synchronisation (ERS) and alpha desynchronisation (ERD), during simulated driving when preceded by a visual search task differing in perceptual load (low versus high) and a passive viewing task. The present study also investigated whether driving performance varied following engagement with the visual search task of low or high load or the passive viewing task. Initial analyses revealed the visual search task successfully manipulated attentional load, as represented by poorer performance, decreased theta synchronisation, and increased alpha desynchronisation during high load search task when compared to low load search task. This replicates the findings from Chapter 3, providing confidence that attentional load was manipulated during visual search. Five main findings related to simulated driving were identified. First, greater parieto-occipital alpha desynchronisation was identified during manual driving when compared to passive viewing. Second, reaction time, lateral control, and longitudinal (speed) control during driving was negatively impacted by the search task, regardless of attentional load. Third, EEG data revealed that frontal and parieto-occipital theta synchronisation during simulated driving was lowest following the passive viewing task when compared to both search tasks. Fourth, parieto-occipital alpha desynchronisation was greatest following the passive viewing task when compared to the high load visual search task. Finally, neurobehavioural correlations revealed a significant negative relationship between parieto-occipital alpha desynchronisation and speed, demonstrating that lower alpha power was associated with better driving performance (closer to the speed limit: 60 mph).

### 4.5.1 Reaction time and standard deviation of lateral position were greater, and speed was lower, following a visual search task of both low and high load

Time to accelerator pedal was greater following a visual search task of both low and high load, when compared to passive viewing, suggesting participants took longer to immediately engage with the driving controls. Takeover quality metrics of lateral and longitudinal control were also negatively impacted by the visual search task. Standard deviation of lateral position (SDLP) was greater following engagement with a search task, when compared to the passive viewing task. This suggests that participants demonstrated less control of the vehicle 'weaving' in and out of the lane. In addition, participants speed was lower following a search task. For the vehicle to keep constant speed, participants were required to accelerate. Therefore, time to accelerator pedal and the lower speed value following a search task represents participants keeping their foot off the pedal and not accelerating when required to. This could be due to allocation of cognitive attentional resources to taking over, and therefore the delayed ability to regain awareness of the vehicle controls such as accelerating. The findings support previous studies that have demonstrated that engagement with prior tasks impact takeover time and takeover quality metrics such as longitudinal acceleration (Gold et al., 2016), speed reduction (Larsson et al., 2014) and lane deviations (Merat et al., 2014). However, there are a number of studies that have contradictory results. Like the present study, Zeeb et al. (2017) found lateral control was impacted by a secondary task. However, braking was not impacted. At first glance, these results seem to differ from the present study, however, Zeeb et al. (2017) measured time to first braking rather than speed deviation from the limit. Therefore, the vehicle speed metric involves the continual control of speed, whereas first braking represents a transient reaction time measure representing an intuitive response to a takeover. Yet, the present study did demonstrate deficits in time to first acceleration, suggesting a transient reaction time measure was negatively impacted by the engagement in an unrelated task, contradicting Zeeb et al's. (2017) findings. Nevertheless, as studies concentrate on distinct tasks and metrics to characterise takeover performance, it is difficult to directly compare experiments.

Although takeover time and quality differed between a passive viewing task and both search tasks, a significant effect was not found between low load and high load search tasks. This is in accordance with previous studies that have similarly found that engagement with tasks impacts takeover performance, regardless of the task performed (e.g. Dogan et al., 2019; Zeeb et al., 2017). However, the results differ from studies that have found differences between type of task. Gold et al. (2015) compared four types of task on takeover performance, each separately involved in different aspects of visual processing, motor processing, and/or cognitive processing. They found that visual processing as measured by the surrogate reference task had an impact on takeover times, yet the cognitive n-back task did not have an impact on easy takeover situations. The authors explain this discrepancy citing the multiple resources model (Wickens, 2002), which is rooted in bottleneck theories of information processing (Broadbent, 1958). The surrogate reference task and manual driving share similar cognitive processes which cannot be carried out concurrently. Therefore, performance deficits are due to the overlap of similar visual processes engaged. On the other hand, in less-cognitively demanding takeovers, engaging in cognitive tasks beforehand has little interference on the takeover. This could explain the present findings. Takeover situations were non-critical nor complex. In addition, participants undertook the same type of takeover 180 times during the one-hour experiment. Therefore, the complexity of the takeover could have been insufficient to evoke a differential effect between low and high load search tasks.

It is challenging inferring takeover quality from limited metrics. Radlmayr et al. (2019) recently introduced a framework for measuring takeover performance including vehicle behaviours such as lateral and longitudinal acceleration, cognitive processing parameters such as gaze reaction times, and subjective ratings such as complexity. While this allows for a more comprehensive way to fully understand takeover performance, the main goal of the present study was to characterise attentional and target encoding neural mechanisms during driving following a secondary task. Nevertheless, the results suggest that lateral and longitudinal control of the vehicle were impacted following engagement with a prior unrelated task.

### 4.5.2 Parieto-occipital alpha desynchronisation was greater during manual driving compared to passive viewing

Additional analyses revealed that parieto-occipital alpha desynchronisation was greater during active driving versus passive viewing, where the participant merely monitored the roadway ahead. Crucially, the visual environment was the same between both conditions, yet the task demands were distinct.

Active driving required recruitment of cognitive and motor processes to enable effective vehicle control, suggesting the increase in alpha ERD represents this increased recruitment and engagement of multiple networks related to successful takeover. Additional analyses revealed that the significant differences were apparent in the first five seconds. During the following 5 seconds, alpha ERD increased during passive viewing, and values were similar to manual driving (Figure 4.10). The experimental design may have encouraged the increase in alpha ERD. During the task, participants were aware that they were required to grab the steering wheel and take control of the vehicle in a few seconds time. An increase in alpha desynchronisation may therefore present this anticipation; participants may have learnt when to start engaging attentional mechanisms for successful takeover. Indeed, the literature reports anticipatory alpha power desynchronisation, particularly when stimulus intervals are predictable (e.g. Foxe et al., 2014; Rohenkohl & Nobre, 2011). Unfortunately, trials of the same length had to be administered for increasing the signal-to-noise ratio of the EEG data. Therefore, it was not possible to have trials of different lengths, with longer intervals and more inter-stimulus variability in time, resulting in less takeovers. Intriguingly, it took approximately 2 seconds for alpha ERD levels to reach similar levels to manual driving, just before the expected takeover was about to happen. Meta-analysis behavioural research has shown that the average takeover time is 2.72 seconds (Zhang et al., 2019), yet participants in the present study were, on average, quicker, as time to accelerator pedal was just over 2 seconds. Did the increase in alpha desynchronisation support a quicker takeover? Although this is speculative, future research should explore this link by comparing alpha levels and takeover times between unexpected and expected takeovers.

### 4.5.3 Frontal and parieto-occipital theta synchronisation was suppressed following a passive viewing task

An increase in theta power occurred at cue onset following both search tasks, yet theta ERS was suppressed following the passive viewing task. This was expected, as phasic theta activity has been suggested to relate positively to task demands (Klimesch, 1999). As participants performed better during simulated driving following the passive viewing task, this suggests that task demands were lower. Moreover, frontal theta has been associated with attentional allocation processes. Missonnier et al. (2006) found that frontal theta ERS was greater during an oddball detection task compared to an n-back task. The detection task required participants to press a button as soon as a target was detected, representing attentional allocation rather than working memory processes. The authors therefore argue that a directed attention network is reflected in the theta response. This is supported by recent research that found increased theta power during externally directed attention (e.g. Kam et al., 2018), as well as research that shows increases in theta ERS during externally directed tasks such as conflict detection (e.g. Töllner et al., 2017). In the present study, switching from passive viewing to manual driving required a lower level of external attention to be initially engaged, due to the same visual environment. Switching from a task to driving required more attention recruiting larger neuronal populations. This could be reflected by increased theta.

### 4.5.4 Parieto-occipital alpha desynchronisation was lowest following the passive viewing task compared to the high load search task

Alpha desynchronisation was greater, measured as lower alpha power, during simulated driving following the passive viewing task, compared to a task of high load. Given that reduced alpha power reflects engagement of the respective region (Klimesch et al., 2007), the results suggest a reduction in task engagement following a high load search task. Although it was hypothesised that lower alpha power following high load (i.e. greater alpha ERD with the change to simulated driving following a search task) due to an increase in attention resource allocation, the results in combination with the behavioural data suggest that deficits in attentional allocation transpired following the high load search task.

Previous research has found that during driving alpha power is greater during self-reported mind wandering (Baldwin et al., 2017) and fatigue (Chuang et al., 2018). Moreover, increased alpha power has been associated with increased driving errors (Campagne et al., 2004). Chuang et al. (2018) investigated oscillatory patterns during one-hour of manual driving. During the study, the vehicle would drift out of the lane, and participants were required to manoeuvre the vehicle back to the central lane position. Alpha ERD was present in the parieto-occipital areas, yet, as reaction times increased, so did alpha power, reflecting reduced ERD. These results are in accordance with the present study, as EEG and behavioural analyses revealed reduced alpha ERD and worse performance following a visual search task. Although following passive viewing task the EEG results show a strengthening of alpha suppression in the parieto-occipital region indicating that participants had to exert more effort to maintain attention as they took over manual driving, task performance was better. Moreover, neurobehavioural correlation analyses revealed that greater alpha ERD was associated with better longitudinal control of the vehicle. This therefore suggests deficit attentional allocation strategies following a high load search task, supported by poorer performance.

Yet, this study did not reveal a significant difference between levels of task load, in that a low load search task and a high load search task had the same effects on parieto-occipital alpha during simulated driving. This is similar to the task performance measures as there was also not a significant difference between search tasks. Potentially, both takeovers were cognitively demanding, although the between task effect size was small. Although the difference between a low load task and the passive viewing task was not significant, the difference between a high load task and a passive viewing task was significant. This therefore suggests that the attentional load of a prior task does modulate takeover neural mechanisms. While a task of low load is no different to watching the roadway (passive viewing task), a task of high load results in reduced alpha ERD. Potentially, the search task may have reduced the immediate re-allocation of attentional resources, as reflected by poorer driving performance and reduced parieto-occipital alpha ERD. The task of high load effectively ‘switched off’ the relevant ERD response.

The correlation between vehicle speed and parieto-occipital alpha was not significant in the passive viewing task condition, yet was significant in both low and high load search task conditions. These results could represent a ceiling effect. As there was little task demand, optimal task performance may have been achieved, or alpha ERD reached maximum resource capacity. As a consequence, the relationship between vehicle speed and parieto-occipital alpha may have been masked. In addition, there was no significant correlation between lateral control of the vehicle and parieto-occipital alpha in any of the conditions. One possible explanation for this is that alpha ERD is involved in specific computations required for longitudinal control. Conversely, alpha ERD is less involved in the computations required for lateral position and more the driving task in general.

In combination, the results suggest that alpha ERD was greatest following the passive viewing task representing successful attentional resource allocation linked to the initiation of manual driving following a takeover. Attention allocation was impaired following a visual search task of high attentional load, represented by lower alpha ERD and poorer performance.

### 4.5.5 Limitations

While the present study offers insight into the oscillatory mechanisms impacted by non-driving related tasks, it is not without limitations. As the aim of this study was to measure event-related synchronisation and desynchronisation, takeovers were repeated approximately every nine seconds. Understanding driving performance over short trials is limiting. In the real world, it is likely that a Level 3-4 autonomous vehicle takeover will occur infrequently. As a result, takeover quality and time may be substantially hindered and therefore unrelatable to a driving simulator study. In addition, repeating specific movements in controlled environments may encourage recalibration, and learning effects have been reported in the handover process (Larsson et al., 2014; Mole et al., 2019). Despite this, a significant difference was detected in behavioural and EEG measures during takeover.

Despite approximately 10% of the population being adextral (e.g. left-handed or showing no preference; Peters et al., 2006), this study, including the studies described in Chapters 2 and 3, included only right-handed participants. This is common practice in cognitive neuroscience research when group averaging approaches are undertaken. If the sample population includes participants with varying cortical lateralisation, then statistical sensitivity will reduce (e.g. motor-related activity over the left and right hemisphere will differ between left- and right-handed participants, adding variance and cancelling out significant differences). Despite this reasoning, it has been argued that ignoring adextral populations hinders our understanding of the diverse human brain and there have been recent calls to include adextral populations in cognitive neuroscience research (Bailey et al., 2019). Although left-handed participants were not strictly excluded from the present study, nor the studies described in Chapters 2 and 3, all participants indicated that they were right-handed for computer mouse use. Therefore, for the case of these studies, they were considered right-handed.

### 4.5.6 Conclusion

In general, previous studies have found that engaging in non-driving related tasks during periods of automation has a significant impact on takeover driving performance, often attributed to deficits in attention and arousal-related mechanisms. Yet, it was not clear how different levels of task demand impacted performance. This study aimed to add to the literature by being the first study to provide neurobehavioural evidence for this deficit. The present study successfully took an established task (visual search) that was used successfully in Chapter 3, to a more applied setting to enable a controlled evaluation of prior task load on simulated driving. Participants switched between non-driving related tasks and simulated driving. Three unrelated tasks were administered: a passive viewing task which involved watching the roadway ahead, a visual search task of low perceptual load, and a visual search task of high perceptual load. Following search tasks, participants performed worse as reflected by increased time to the accelerator pedal, greater speed deviation from the speed limit, and increased standard deviation of the lateral position, reflecting poorer vehicle control. A transient burst of theta synchronisation was absent following the passive viewing task, reflecting fewer encoding mechanisms potentially due to the similar visual environment. Parieto-occipital alpha ERD activity was less pronounced following the search task of high load when compared to the passive viewing task, representing deficits in attentional resource allocation. Furthermore, alpha power negatively correlated with speed, implying a direct link between performance and neural attention-related activity. Overall, the study provides the first behavioural and neural evidence that engagement with non-driving related tasks of high attentional load reduces the extent of processing of a simulated driving task. These results have significant implications for future research and the implementation of semi-autonomous driving. If a driver is engaged with a task of high attentional load, and a takeover occurs in an emergency situation, the driver may not be able to recruit the appropriate attentional mechanisms quickly enough to safely regain control of the vehicle, potentially having life-threatening consequences. In addition, the results highlight important neural mechanisms for successful takeover performance, particularly parieto-occipital alpha activity. This provides potential insight into the possibility for user training in monitoring of automated systems, such as drivers in semi-autonomous vehicles, or pilots in airplanes. Using neurofeedback, introduced in Section 1.3.2, users can learn when to increase/reduce neural activity to improve human performance. In the case of this Chapter, drivers might be able to learn how to reduce alpha power to improve driving performance following a takeover.

Although this study provides neural evidence for attention deficits following unrelated tasks, it is difficult to undertake accurate and reliable EEG measurements in the real world. Therefore, the following study in the next Chapter explores the feasibility of measuring attentional deficits as indicated by physiological signals unobtrusively, and how they are represented during real-world autonomous driving. This should provide further insight into the possibility for utilising a biofeedback system to engage attention in naturalistic environments.

# 5.0 The Impact of an Unexpected Event on Autonomic Arousal and Eye Fixations during Autonomous Driving

## 5.1 Overview

The previous Chapters revealed specific attention-related neural oscillatory patterns and performance deficits during a visuomotor task following unrelated visual search tasks differing in attentional load. The findings from these Chapters have a wide range of implications indicating that changes in attention are likely when switching between tasks. During semi-autonomous driving, for example, passive drivers will have the opportunity to engage in non-driving related tasks, and completely disengage with the vehicle and driving environment, although should continue to monitor the automation for any faults. Similarly, safety drivers responsible for an automated vehicle under development can disengage from monitoring the automation, despite the catastrophic consequences following from overlooking automation errors and potential hazards in the driving environment. However, applied research is needed to explore the feasibility of measuring fluctuations in attention and how fluctuations of attention are represented, during real-world autonomous driving. In addition, given that human-vehicle collaboration and physiological state depends on an array of factors including age, a comprehensive understanding of specific populations of participants are needed to better inform a driver state monitoring system. Considering the potential autonomous vehicle benefits for older adults, such as maintaining mobility and independence, the present study focused on older adults' physiological state during periods of automated driving. During this study, multi-modal attention measures including subjective ratings, autonomic arousal, and eye gaze, were collected. To fluctuate arousal levels during autonomous driving, participants experienced two types of stops during different journeys. During an unexpected hazardous stop, a pedestrian walked in front of the vehicle causing the vehicle to brake. During an expected stop, the vehicle stopped due to journey route set up.

## 5.2 Introduction

Road traffic accidents are amongst the top 10 causes of mortality worldwide and are the leading killer of people aged 5 to 29 years (World Health Organisation, 2018). As vehicles move from manual to semi-autonomous to autonomous technology, accident rates and fatalities should reduce as human errors and judgement are likely to be removed from autonomous driving situations (Fagnant & Kockelman, 2015). Yet, research on automation has demonstrated that rather than the human being eliminated, the role of the human changes: humans must adapt to the role of monitoring automation, which often leads to negative consequences associated with the misuse and disuse of automated systems (Parasuraman & Riley, 1997). Misuse refers to when the user relies on automation when it performs poorly; whereas disuse refers to when the users fail to engage automation when it will improve performance.

### 5.2.1 Allocation of attention during automated driving

Misuse may play a significant role in the safe development of highly automated vehicles as periods of automated driving will have a strong impact on the allocation of drivers' attentional resources. Typical performance indicators of attentional allocation are not appropriate during autonomous driving, as the passive driver is not required to carry out manual driving behaviours. Therefore, measures of speed or lateral position changes will not provide any insight into human performance, as the automation is in control. Other methods able to detect attentional allocation must be used. Attentional deployment depends on time-varying factors, and therefore retrospective ratings cannot truly capture attention levels during complex perceptual and cognitive experiences. Capturing the human response in real-time can provide light on the human experience during dynamic driving scenarios. Continual measures of eye gaze and physiology can provide a deeper understanding of transient and sustained attention during sudden events and long-lasting driving conditions.

Gaze has been shown to play an important role in the driving task with 80% of driving information obtained by eye movements (Kowler, 2011). Visual strategy and the distribution of fixation points can provide information about where and when participants are shifting their attention, measured via fixation count and fixation duration, respectively. Therefore, most studies have investigated visual gaze behaviour during periods of partially or fully automated driving. Several studies have attempted to understand the associations between constructs related to automation monitoring and attention itself. For example, participants with a high level of trust tend to monitor the road less (Helldin et al., 2013; Hergeth et al., 2016; Körber et al., 2018; Walker et al., 2019). Situation awareness has also been associated with gaze behaviour. Shinohara et al. (2017) found that recall of features in the driving environment increased with fixation count and duration, suggesting that participants' situation awareness was greater with efficient visual gaze behaviours. Research has also indicated that psychological constructs such as trust and situation awareness are intermixed. Petersen et al. (2019) found that enhancing situation awareness by providing direct information regarding driving execution increased levels of trust.

The aforementioned studies were undertaken in simulators, yet real autonomous driving is likely to place a greater demand upon the sensory and cognitive systems. For example, research has found that drivers tend to steer in the direction of their gaze (Robertshaw & Wilkie, 2008), and as attentional demands increase during real-world driving, visual scanning behaviours reduce (Recarte & Nunes, 2003; Savage et al., 2013). One study attempted to understand attentional state in passive drivers by investigating the impact of real-world driving on passengers. Takeda et al.'s (2016) study assumed that a passive driver is similar to a passenger. Their results revealed that large saccades were greater in passengers when compared to drivers, suggesting that passengers were more likely to scan irrelevant stimuli. Eye-blink duration was also greater in passengers, indicating lower arousal. In combination, these results indicate that attentional load was lower in passengers. However, the study fails to capture the true role of a passive driver in an autonomous vehicle, whom will have access to a type of human-machine interface for active monitoring of the automated system, and may be able to override the automated vehicle controls by activating an emergency stop button.

Suboptimal levels of cognitive functioning can also be assessed via psychophysiological measures of autonomic arousal (Lohani et al., 2019). Carsten et al. (2012) found that heart rate was lower during autonomous driving when compared to semi-automated and manual driving, providing support for cognitive underload during periods of automation. These results are supported by studies that have found that autonomous driving reduces self-reported workload (De Winter et al., 2014). The serious implications of cognitive underload during autonomous driving have been emphasised by studies investigating sleepiness and fatigue. Vogelpohl et al. (2019) found that passive drivers in automated driving situations reported greater fatigue levels and displayed prolonged eyelid closures, reflecting lower arousal or sleepiness when compared to manual drivers.

Neurophysiological evidence provides additional support that attentional mechanisms are modulated during periods of automated driving. Van Der Heiden et al. (2018) found the P3a event-related potential component, related to novel stimuli, was reduced in amplitude during automated driving when compared to stationary. These results indicate that during periods of automation, the processing of task-irrelevant auditory stimuli is reduced. Despite Takeda et al.’s (2016) study indicating that passengers scan task-irrelevant stimuli more than manual drivers, the results from Van Der Heiden et al. (2018) suggest that cortical processing of irrelevant stimuli is lower in passive drivers. These conclusions are partially supported by Hidalgo-Muñoz et al.'s (2019) work measuring oxyhaemoglobin with functional near-infrared spectroscopy. Oxyhaemoglobin was lower, reflecting increased neural metabolism and activity, in right frontal areas during autonomous driving, compared to manual, when listening to auditory stimuli. These results suggest that during automated driving, attentional resources were allocated to processing irrelevant stimuli more so than when manual driving. In line with perceptual load theory (Lavie, 2005, 2010), the more difficult the task (i.e. driving versus non-driving), the more difficult it is to process background stimuli. Therefore, during periods of automated driving, attentional mechanisms are more susceptible to distractors.

### 5.2.2 Allocation of attention during unexpected events while driving

Overall, the current research suggests that information processing load is lower during autonomous driving. However, attentional allocation evolves over time with changing task demands. Therefore, understanding how attention is modulated during context-specific scenarios can provide insight into the mechanisms responsible for insufficient responses when automation fails. Measures of visual behaviour have demonstrated distinct oculomotor actions during unexpected events. During normal driving, fixation duration increases and visual scanning decreases during hazardous moments (Mackenzie & Harris, 2015). Strauch et al. (2019) found that participants fixated in safety-critical areas (i.e. the steering wheel and forward roadway) more so during automated versus manual driving. This suggests that critical events during autonomous driving are more demanding on attentional systems, thus emphasising the need for a sufficient amount of available attentional resources.

Increases in arousal have been linked to attention narrowing to the central focus of stimuli (e.g. Laumann et al., 2003). Therefore, the breadth of attentional focus can also be demonstrated by physiological indicators of arousal during unexpected events in automated driving (Meinlschmidt et al., 2018). For example, Zheng et al. (2015) found that masseter electromyography increased and self-reported comfort decreased, as the headway between the lead vehicle decreased. During unexpected takeover requests and misleading notifications, Ruscio et al. (2017) demonstrated an increase in sympathetic arousal as measured by skin conductance response amplitude. In contrast, parasympathetic activity, as measured by respiratory sinus arrhythmia, did not change. However, parasympathetic inhibition did increase during expected takeover requests. Analogous to the malleable attentional resources theory (Young & Stanton, 2002), the authors argue that attentional capacity was hindered during periods of automation. An unexpected takeover generated an overload of central processing resources (i.e. no parasympathetic inhibition), whereas an expected takeover did not. Therefore, during the unexpected takeover, attentional resource allocation was reduced, leading to delayed reaction times.

Overall, the research demonstrates that during an unexpected critical event, attentional demands are disproportionate to the allocated resources. Autonomous driving encourages attentional demands to become even more inadequate due to factors such as fatigue, loss of situation awareness, overtrust and overreliance (e.g. Choi & Ji, 2015; Vogelpohl et al., 2019). In turn, this could lead to unsafe human-vehicle interaction and subsequently critical accidents.

### 5.2.3 Older adults and autonomous driving

Attentional allocation and automation issues discussed in the previous sections are potentially amplified in an older adult population with ageing-related impairments, as they are more likely to rely on automated systems (McBride et al., 2011), find it more difficult to perform two or more tasks simultaneously (Kramer & Madden, 2008), and are more prone to lack understanding of advanced technology (Mann et al., 2007). Moreover, research has indicated that older adults have concerns using autonomous vehicles due to issues related to trust and confidence, such as not having an operator nearby during autonomous vehicle failures (Faber & van Lierop, 2020). Yet, autonomous driving has the potential to enhance older adults' wellbeing and health. Age-related declines in cognitive, visual capacities, physical disability, and illness, subsequently impact driving ability as it becomes more physically and cognitively demanding. The possibility of becoming a non-driver rises with age (Anstey et al., 2006), and some drivers choose to restrict their driving (Dellinger et al., 2001). Driving cessation can have a negative impact on mobility and wellbeing, and feelings of isolation can be amplified (Qin et al., 2019). As different cognitive and physical demands of the task are replaced by automation elements, autonomous vehicles may offer an alternative transport solution for the older population. By enabling a viable transportation option, mobility is likely to be restored enabling older adults to lead more independent lives. In turn, this should promote participation in local and social events, encouraging feelings of social inclusion and satisfaction.

Considering the potential autonomous vehicle benefits for older adults, such as maintaining mobility and independence, and the age-related individual differences related to human-automation collaboration, a comprehensive understanding of older adults' psychophysiological state during periods of automated driving, particularly during unexpected situations, is needed. Understanding an older adults' functional state will better inform a driver state monitoring (DSM) system including cognitive and affective indices, to improve health and wellbeing for older adults travelling in autonomous vehicles. This is particularly important as autonomous vehicles have the potential to enrich older adults' lives, but only if they are implemented successfully.

### 5.2.4 Experiment rationale

The goal of the present study was to understand how visual attention and arousal fluctuated in older adults during an unexpected event during autonomous driving. An experiment was designed which instigated an unexpected stop due to a hazard during real-world autonomous journeys. Subjective measures of trust, reliability, workload, and situation awareness were taken retrospectively, before and after all autonomous journeys. Trust and reliability were also rated after the unexpected stop. Physiological measures of electrodermal activity and heart rate, and gaze fixation metrics were collected continuously throughout the experiment, to provide further insight into arousal mechanisms, and transient and sustained attention. The unexpected stop was compared to an expected stop.

This experiment was an applied field study using a real-world autonomous 'Pod'. Although a number of studies have suggested that simulators can provide a valid tool for assessing driving behaviour (Mullen et al., 2011), it is not clear whether simulators can assess drivers' attentional allocation during autonomous driving. Therefore, studies utilising real vehicles are necessary to understand the impact of unexpected events on drivers' attention. In addition, several subjective, physiological, and cognitive indices were collected to try and accurately capture human-vehicle interaction during the unexpected stop. Chapter 2 revealed an increase in skin conductance level during the first minute of a task, indicating a transient autonomic response to a demanding situation. In addition, heart rate increased halfway through the task, indicating a slower autonomic response to a demanding task. Therefore, electrodermal activity in combination with heart rate was collected for the present study. Ocular behaviour provided real-time measures of transient and sustained attention, indicating which areas of the visual environment captured eye gaze, and how visually and cognitively demanding the visual environment was. EEG indices were not collected for several reasons. Firstly, set up of the EEG, which normally takes up to an hour, would have increased the duration of the study drastically, resulting in fatigue and boredom confounding potential findings. Secondly, it would have been difficult to set up an EEG in the small pod-like vehicle that moved around freely, as the equipment is not portable. A wireless dry sensor EEG system could alleviate some concerns, but would still require some preparation, and advanced technical expertise to send event markers (i.e. the onset of an unexpected event), as well as restricted functionality due to low sampling rate and susceptibility to artefacts such as facial muscles. In addition, an event-related design could not be carried out and the signal-to-noise ratio of the EEG would have been reduced. Ultimately, given the applied nature of the study and the sample population, less obtrusive and portable devices were preferable for participant comfort and acceptability.

The hypotheses were informed by previous research suggesting attentional demands are disproportionate to the allocated resources during unexpected events while driving (e.g. Ruscio et al., 2017; Strauch et al., 2019; Zheng et al., 2015). Therefore, it was predicted that the unexpected event would narrow the focus of overt visual attention coupled with an increase in heart rate and skin conductance and a reduction in heart rate variability. As autonomous driving has been shown to induce fatigue, reduce situation awareness, and promote overtrust and overreliance in older adults (e.g. Choi & Ji, 2015; Vogelpohl et al., 2019), it was expected that self-report ratings of trust and reliability would be greater following an unexpected event compared to an expected event, and situation awareness and workload ratings would be reduced following autonomous driving.

## 5.3 Method

### 5.3.1 Participants

Thirty-nine adults originally participated in this study. Two participants were excluded from all subsequent analyses due to the Pod experiencing technical errors during the journeys, leaving 37 participants (sixteen females, twenty-one males, mean age ± SD = 68.35 ± 8.49 years, range 48 – 89 years, two participants under 60 years). Due to recording errors during data collection, only 30 participants' physiological data were subsequently analysed (twelve females, eighteen males, mean age ± SD = 69 ± 8.75 years, range 48 – 89 years). Due to eye-movement abnormalities such as lazy eye (three), technical errors including unsuccessful calibration of the eye tracker (five), and low gaze samples (three), only 26 participants' eye tracking data were subsequently analysed (twelve females, sixteen males, mean age ± SD = 67.19 ± 7.32 years, range 52 – 89 years).

All participants had normal or corrected-to-normal vision. Five participants had corrected hearing. All but three participants held a valid driving license. No participants had any previous experience with highly automated driving. Those with significant health conditions (e.g. epilepsy, neurological impairments, coronary issues) were not permitted to take part. Participants received a £20 voucher as compensation for their participation to cover expenses. All participants gave written informed consent and were fully debriefed at the end of the study. Ethical approval was obtained by the Faculty of Health and Applied Sciences University of the West of England Research Ethics Committee (HAS.18.09.024).

### 5.3.2 Apparatus

#### *5.3.2.1 Autonomous vehicle*

A Pod Zero autonomous pod provided by Aurrigo (RDM Group) was used as the autonomous vehicle (see Figure 5.1). The Pod is a compact research and development vehicle designed to be used in pedestrian areas and shared pedestrian/vehicle routes. It is electrically driven and can be used continuously for 10+ hours of normal operation. It is a four-seater vehicle, with two benches facing each other designed similarly to a four-seater in a train. Due to safety regulations, a safety person was always present in the vehicle observing the environment and had access to an emergency stop button. The participant sat facing forward, viewing the external environment out of a front and two small side windows (see Figure 5.3 for an example of a participants' field of view). Four marshals supervised the front and back of the vehicle, and the route was supervised by additional marshals at each intersection to ensure no vehicles or pedestrians caused an obstruction.

The autonomous behaviour of the Pod was achieved using the Wizard of Oz approach (Kelley, 1985), whereby the Pod was remotely teleoperated in manual mode using a hand-held wireless control unit by an operator positioned behind the vehicle not in view of the participant. Driving the Pod in the teleoperated mode ensured that its actions were replicable between participants; the Pod could be made to respond similarly to different obstacles and follow the route as planned. At the beginning of the study, participants were told the vehicle was run fully autonomously. During debriefing, participants were told the Pod was operated manually by a teleoperator walking behind and remotely controlling it during the study. As the driving route involved a pedestrian area, the vehicle was controlled at walking speed, approximately 3-5 mph.

#### *5.3.2.2 Human-machine interface (HMI)*

The human-machine interface (HMI) was presented on a HannsG HT161HNB 15.6" Multi Touch Screen connected to a Kodlix GN41 Mini PC (Windows 10, Intel Celeron processor, 8GB RAM, 64GB). The design of the HMI was informed by HMI design principles, public engagement workshops with older adults, and feedback from previous iterations of the HMI (Eimontaite et al., 2020; Morgan et al., 2018).

The HMI graphical touch screen displayed the vehicle speed, time remaining until destination, a safe stop button, a journey map, vehicle ‘health’, and journey set up/change options (see Figure 5.1). The functionality of the safe stop button was described to the participant at the beginning of the study, emphasising that pressing this icon would initiate the vehicle to stop. The vehicle ‘health’ icon provided information about the current working order of the automated system including the tyres, brakes, network, and battery level. During the study, the vehicle health was always shown as being in good working order. The HMI presented visual and audio notifications to describe the vehicle’s behaviour and journey course, such as “*Turning left*” and “*You have arrived at your destination*”.

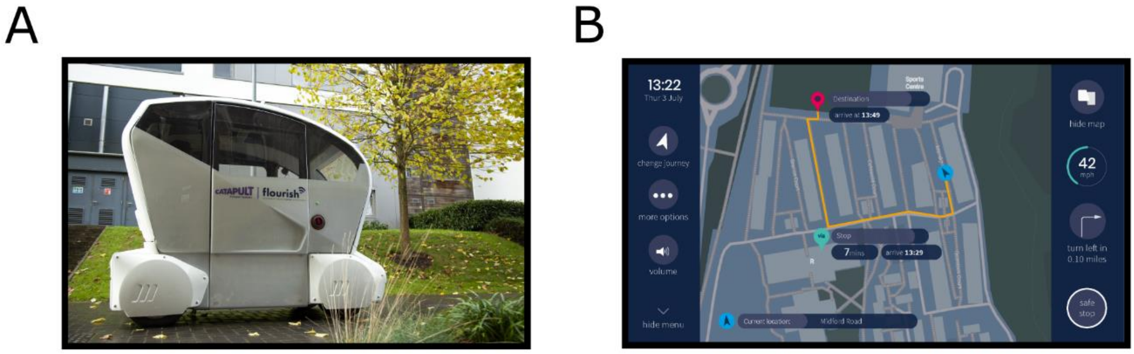



*Figure 5.1.* Autonomous vehicle and human-machine interface. **(A)** Autonomous Pod utilised during the study. **(B)** Human-machine interface display during the journey.

### 5.3.3 Journeys

Participants in the study went on six journeys consecutively in a random order. Before each journey, participants were provided with a scenario that specified the journey they were required to set up on the HMI. There were six possible destinations/stops in total: Home, Health Centre, Recycling Centre, Sports Centre, Sports Field, Post Office. Among the six journeys there was always a journey including an expected stop and another journey including an unexpected stop due to the hazard. Both journeys were of an equivalent length and lasted for approximately six minutes. Some of the other journeys also included other variables such as picking up a friend. As the main focus of the Chapter is to investigate the impact of an unexpected event, other journeys will not be described in detail. All journeys were randomised between participants so that the unexpected stop happened either during journey two or journey five, and the expected stop happened during journey one, journey three, or journey four. Overall, participants experienced approximately 60 minutes of the automated driving system.

The expected stop was initiated during journey set-up and was therefore expected by the participants. A few seconds before the vehicle stopped, an HMI notification "*You are arriving at [Stop]*" was presented. Once the vehicle stopped, a notification "*You have arrived at [Stop]*" was presented. The HMI then displayed an option to either resume or stop the journey. All participants resumed the journey. The unexpected stop was executed as an emergency stop appearing to the participants as happening suddenly, and as such, was not anticipated by the participants. A marshal was instructed to answer their mobile phone and walk in front of the Pod. The teleoperator of the Pod then initiated the vehicle to stop. The HMI notification "*The vehicle detected a hazard in the road and has stopped. Your journey will resume shortly*" was presented on the HMI. Once the marshal had moved safely out the way, the Pod would restart and continue the journey. The participant was not required to do anything.

### 5.3.4 Protocol

Participants arrived and met the author and other researchers near the student accommodation area on the university campus, where the trial took place. Participants were reminded of the content of the information sheet, asked about their wellbeing and whether they had any concerns or questions. Then they signed printed copies of the consent form and filled in the pre-trial questionnaires. Once the physiological and eye tracking equipment were set up, participants were taken outside and introduced to the Pod. Participants sat inside the vehicle wearing a seatbelt and facing forwards. At the beginning of each journey, the participant received the journey scenario that specified the journey destination and stop if there was one. Once the journey was set up, the participant experienced the journey. Participants were told they could interact with the HMI as little or as much as they wished to. After each journey, participants provided verbal trust and reliability ratings to the author or another researcher. This process was repeated six times and all participants completed six journeys. Afterwards, participants left the vehicle and filled in several post-trial questionnaires. The full testing session, including the induction and filling out questionnaires, lasted for approximately 150 minutes, depending on inter-individual variability. See Figure 5.2 for a schematic of the experimental procedure. A significant amount of time was required to be scheduled when conducting studies with the sample of older participants. It was important to ensure a pace that did not increase fatigue, and enough time to reflect and discuss issues and questions raised.

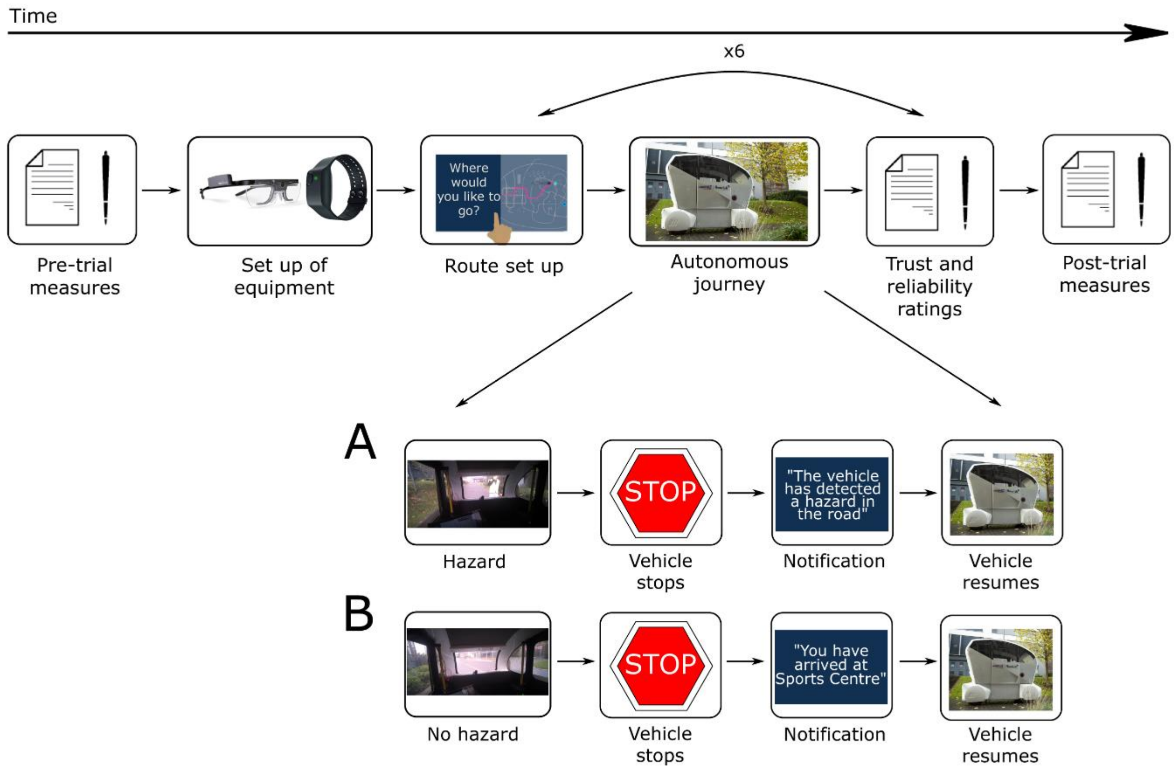



*Figure 5.2.* Experimental procedure. **(A)** During an unexpected event, a marshal walked in front of the vehicle. This caused the vehicle to stop and display a message, *"The vehicle has detected a hazard in the road. The vehicle will resume shortly".* The vehicle resumed once the roadway was clear. **(B)** During an expected event, the vehicle came to a stop when it reached a destination. The HMI displayed a message, for example, *"You have arrived at Sports Centre".* The participant was required to press *"Resume"* on the HMI for the vehicle to resume the journey.

### 5.3.5 Measures

#### *5.3.5.1 Self-report measures*

The trial involved a combination of questionnaires to fill in before and after all (six) autonomous vehicle journeys. The below subjective rating questionnaires were analysed for this Chapter.

*The NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988)*: The NASA-TLX is a subjective, multidimensional questionnaire that assesses mental, physical, and temporal demand, effort, performance, and frustration constructs. The NASA-TLX has been used widely in the manual driving and automated driving literature (e.g. Chen et al., 2019; Saffarian et al., 2012). A description for each of the subscales was provided to the participant before rating. Participants rated each factor within 10-points with .5-point steps. The greater the score, the more demanding was the factor. See Appendix 5.1 for an example of the questionnaire.

*The Situation Awareness Rating Technique (SART; Selcon & Taylor, 1990):* The SART is a multidimensional scaling technique of perceived situation awareness. Situation awareness is a construct that refers to the perception of environmental cues, with a focus on the near past, current, and future situations (Flach, 1995). For example, when a passive driver perceives and detects objects in the driving environment and understands the vehicle's state and behaviour, their situation awareness is considered to be high. Yet, when they are unaware of objects in the visual environment and unaware of the vehicle's state or behaviour, their situation awareness is low. Transportation incidents have indicated that poor situation awareness during automation monitoring can lead to catastrophic accidents (Endsley, 1996), and the SART has previously been utilised in autonomous vehicle research (e.g. Petersen et al., 2019; Schewe et al., 2019). Ten items were presented with bipolar responses e.g. *How changeable is the situation? Is the situation highly unstable and likely to change suddenly (High) or is it very stable and straightforward (Low)?.* Participants responded on a scale from 1 "High" to 7 "Low". See Appendix 5.2 for an example of the questionnaire.

*Trust in Automation (TiA; Gold et al., 2015):* The TiA questionnaire has been used in automated vehicle research to understand the confidence participants' hold in the automated system and the factors that modulate this relationship (Gold et al., 2015). The scale contains 19 items on a Likert-type rating scale from 1 "Strongly disagree" to 5 "Strongly agree" that measure participants' trust in automation e.g. *The system state was always clear to me*. The questionnaire is structured into five subscales: reliability/competence, familiarity, trust, understanding, and intention of developers. See Appendix 5.3 for an example of the questionnaire.

*Trust and reliability ratings:* Trust and reliability were measured with a single-item scale to limit interruptions to the autonomous vehicle journeys. Participants were asked to rate how much they trusted the autonomous vehicle on a scale from 0 "Did not trust" to 10 "Completely trust". They were also asked how reliable the vehicle was on a scale from 0 "Not reliable" to 10 "Completely reliable". In addition to pre-journey and post-journey ratings, trust and reliability scores were collected verbally from participants at the end of every journey.

*5.3.5.2 Physiological arousal*

Continuous physiological acquisition of heart rate (beats per minute; BPM) and electrodermal activity (skin conductance level; µS) were collected using an Empatica E4 wristband (Empatica Inc., Cambridge, MA, USA and Milan, Italy) to measure levels of autonomic arousal. The sampling frequency for the electrodermal activity sensor was 4 Hz and the photoplethysmography sensor on the Empatica measured blood volume pulse at 64 Hz. The internal Empatica software derived the BPM. The Empatica E4 wristband was placed on participants' non-dominant wrist to reduce the possibility of motion artefacts. The Empatica was fastened tightly as comfortable for the participant, so the wristband did not move around. The E4 also collected acceleration data from a 3-axis accelerometer, which enabled monitoring of wrist movements. The sampling frequency of the accelerometer was 32 Hz.

An event marking button on the Empatica E4 was pressed in front of a camera, which triggered a LED light to be illuminated on the Empatica, and simultaneously logged a timestamp in the data. This mode of creating a marker was done to aid the later analysis of when events of interest (i.e. the unexpected event) occurred in the physiological data. The process for deriving the times of interest is described in Section 5.3.6.2.

*5.3.5.3 Eye tracking*

Tobii Pro Glasses 2, an eye tracking device, was used to collect fixation metrics (Tobii Glasses Eye Tracker, Tobii Technology, Stockholm, Sweden). The Tobii Glasses are a wearable eye tracker worn like a pair of glasses. The design is lightweight and has no side or bottom frame, preventing any distraction in the participant's visual field. The head unit is comprised of several cameras; a high-definition camera captured participants' field of view (82° horizontal and 52° vertical), and two eye tracking sensors below each eye captured participants' pupil diameter and movements. To improve the accuracy of the eye tracking sensors, near infrared lights illuminated the pupil. The sensors have a sampling rate of 100 Hz. A microphone recorded the audio of the environment. The Tobi Pro Glasses do not work with standard eyeglasses as glasses can create additional glint that can lead to data corruption. Individuals wearing glasses were asked to remove them, and a suitable prescription lens was attached to the glasses (ranging from –5.0 dpt to +3.0 dpt). Once the participant was wearing the head unit, the manufacturer's calibration procedure was followed which typically took less than 30 seconds and consisted of participants fixating on a central target.

### 5.3.6 Pre-processing

#### *5.3.6.1 Self-report measures*

All questionnaire data were entered into Microsoft Excel 2016 for pre-processing. The NASA-TLX constructs were mapped to a 100-point scale by counting the number of lines the participant marked, subtracting one, then multiplying by five. The constructs were then averaged independently to provide a raw-TLX score for each factor (mental, physical, temporal, effort, performance, frustration). This approach has been widely used in the literature and was undertaken to better understand what aspects of workload were affected (Hart, 2006). An overall SART score was derived using the following formula as provided by the authors (Taylor & Selcon, 1990): summed understanding divided by (summed attentional demand – summed attentional supply). An overall score for TiA was derived from the average of all items.

#### *5.3.6.2 Physiological arousal*

Data were opened and pre-processed in Microsoft Excel 2016 using Excel's in-built functions. Electrodermal activity and heart rate values, with corresponding timestamps, were pre-processed separately and followed the same procedure. For the EDA data (4 Hz sampling rate), every four samples were averaged to produce one value for every second, and similarly, one second averages were used to analyse heart rate data. The averaged data were aligned to the appropriate time point, to allow for averaging across time points of interest. To do this, the timings of the conditions were extracted from the eye tracking recordings. The wearable eye tracking device had a wide angled camera that captured the experiment from the participant's view, and each video sample specified the 24-hr time. The time stamp of the start of the recording was extracted (e.g. 13:07:13). At each time of interest, the 24-hr time was calculated (e.g. Journey 1 started at 05:46 into the recording, and therefore started at 13:12:59). The time the marker was pressed in the recording was subtracted from the 24-hr time of interest, which provided a difference time between the marker and start of the time interest (e.g. 15 seconds). The Empatica marker on the Empatica recording was then converted from UNIX to the corresponding 24-hr time. The Empatica marker time was then added to the difference time. This provided an initial start time stamp for the time of interest in the Empatica data. Next, the appropriate data values were extracted for each time of interest. During pre-processing, it became clear that there were many missing interbeat interval samples from the Empatica. Unlike the blood volume pulse signal which has a sampling frequency of 64 Hz, the interbeat interval signal is only calculated when a beat is detected. Missing values impact the reliability and accuracy of the calculation of the differences between successive samples (Gruden et al., 2019), and therefore, heart rate variability measures of RMSSD, SDSD, and pNN50 were not further analysed. For data relating to the unexpected and expected stop, data were averaged within two times of interest: 30 seconds before the stop and 30 seconds after the stop.

#### *5.3.6.3 Eye tracking*

Eye tracking analysis was undertaken using Tobii Pro Lab software version 1.138 (Tobii Technology, Stockholm, Sweden). The gaze sample percentage across the entire recording was first assessed. Eye tracking glasses captured a mean of 80% (*SD* = 18%) of gaze samples. Events were then logged to indicate the start and end of events in the recording. Times of Interests (TOIs) were defined by selecting the appropriate start and end event markers. This allowed for segmentation of the data into intervals of time relevant to subsequent data analysis. The ‘Pre-stop’ TOI was considered the 30 seconds before the presentation of the notification when the vehicle stopped; the ‘During’ TOI consisted of the time the notification was displayed visually; the ‘Post-stop’ TOI was considered the 30 seconds after the presentation of the notification. Gaze data from the recording were then manually mapped onto an image best depicting the overall visual view of the participant. Next, Areas of Interests (AOIs) were defined on each mapped image for each TOI (see Figure 5.3). Three AOIs were created representing the HMI, the central view of the driving environment, and the peripheral view of the driving environment. To finish, the I-VT Filter (Fixation) was applied to the data, which set the velocity threshold parameter at 30 degrees/second. If the sample was below this threshold, it was classified as a fixation.

Total fixation duration and total fixation count metrics were exported. Because the time of the critical event varied across participants, and to standardise to account for variability within patterns of fixations, it was necessary to calculate fixation count and fixation duration proportions based on the total numbers of fixations and fixation durations. Fixation duration was defined as the amount of time spent looking at each AOI divided by the total duration of fixations. Fixation count was defined as the number of fixations towards each AOI divided by the total number of fixations. Averages were calculated for subsequent analyses.

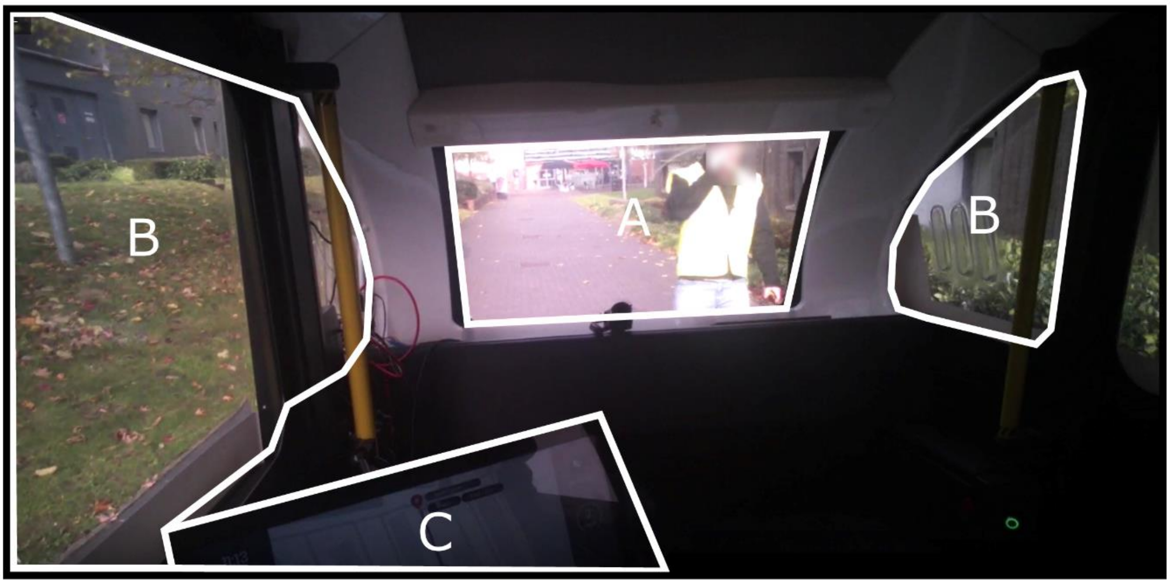



*Figure 5.3*. Areas of Interest (AOI) for eye tracking analysis. **(A)** Central environment. **(B)** Peripheral environment. **(C)** Human-machine interface.

### 5.3.7 Statistical analyses

All statistical analyses were performed using IBM SPSS Statistics for Windows version 26 (IBM Corp., Armonk, N.Y., USA). Descriptive statistics were performed, and normality was verified using the Shapiro-Wilk test and visualization of QQ plots of the unstandardized residuals. Give the applied nature of the research, no data were collected during true baseline or recovery periods, and so z-scores were calculated to standardise the heart rate and electrodermal activity data due to the individual variability of physiological responses (Braithwaite et al., 2012). Assumptions of sphericity were tested using Mauchly's test and, if violated, Greenhouse-Geisser estimates were used in the repeated measures calculations. The statistical threshold for significance was set to two-tailed $p < .05$. Effect size was reported as eta squared ($\eta^2$) for one-way ANOVA significant results and partial eta squared ($\eta_p^2$) for two-way ANOVA significant results (Cohen, 1988). Cohen's $d_z$ was reported for paired-samples *t*-tests. *Post hoc* analyses were run with Bonferroni correction.

*Self-report ratings of workload, situation awareness, trust in automation, trust and reliability:* For workload (NASA-TLX), a 6 (Factor: mental, physical, temporal, effort, performance, frustration) × 2 (Time: pre-, post-) repeated measures ANOVA model was run. Paired-samples *t*-tests were used to compare pre-journey and post-journey measures: situation awareness (SART) and trust (TiA). For single-item trust and reliability ratings, a one-way repeated measures ANOVA (Journey: pre-, unexpected, expected, post-) was undertaken.

*Physiological arousal:* A 2 (Stop: unexpected, expected) x 2 (TOI: 30 s before, 30 s after) repeated measures ANOVA was performed to understand the impact of an expected and unexpected stop on physiological arousal. The model was run for both heart rate and skin conductance level *z*-scores.

*Eye tracking:* Two two-way repeated measures ANOVA were undertaken on both fixation count and fixation duration measures. The first was a 2 (Stop: unexpected, expected) x 3 (AOI: central, peripheral, HMI) repeated measures ANOVA to understand the impact of journey type on AOI. The second ANOVA was a 2 (Stop: unexpected, expected) x 3 (TOI: pre-stop, during, post-stop) repeated measures ANOVA to understand the impact of journey type on time.

## 5.4 Results

### 5.4.1 Self-report measures

*Workload (NASA-TLX):* The repeated measures ANOVA model showed the main effect of Time was not significant, $F_{(1, 36)} = 1.36$, $p = .25$. The main effect of Factor was significant, $F_{(3.67, 131.96)} = 35.53$, $p < 0.001$, $\eta_p^2 = 0.50$, as was the interaction Time by Factor, $F_{(3.98, 143.24)} = 6.06$, $p < .001$, $\eta_p^2 = 0.14$. Post-hoc comparisons revealed that Physical ($p < .001$), Effort ($p = .039$), and Frustration ($p < .001$) workload factors differed between pre- and post- measures. While Physical and Effort factors reduced after the autonomous journeys, frustration increased following autonomous journeys. See Table 5.1 for descriptive statistics.

*Situation awareness (SART):* A paired-samples *t*-test revealed that situation awareness was similar between pre- and post- journeys, $t_{(36)} = 0.39$, $p = .675$. See Table 5.1 for descriptive statistics.

*Trust in automation (TiA):* A paired-samples *t*-test revealed that trust in automation increased after the autonomous journeys, when compared to before the autonomous journeys, $t_{(36)} = -2.95$, $p = .006$, $d_z = -0.48$. See Table 5.1 for descriptive statistics.

*Table 5.1.* Mean (SD) of subjective ratings for situation awareness, workload, and trust in automation before and after autonomous vehicle journeys.

| Subjective rating | Journeys | |
|---|---|---|
| | Pre- | Post- |
| **SART** | | |
| Situation awareness | 22.15 (11.72) | 21.41 (7.02) |
| **TiA [0-5]** | | |
| Trust in automation* | 3.39 (0.33) | 3.57 (0.33) |
| **NASA-TLX [0-100]** | | |
| Mental | 23.65 (13.43) | 22.70 (11.79) |
| Physical** | 14.60 (11.23) | 8.85 (9.06) |
| Temporal | 16.48 (12.10) | 16.35 (12.13) |
| Effort* | 17.16 (12.75) | 11.82 (11.56) |
| Performance | 21.62 (12.71) | 18.89 (11.31) |
| Frustration** | 5.41 (4.95) | 10.51 (10.38) |

*Key:* **p* < 0.05; ** *p* < .001

*Single-item trust and reliability ratings:* For trust ratings, a one-way repeated measures ANOVA model revealed a significant main effect for Journey, $F_{(2.05,\ 73.86)} = 15.05$, $p < .001$, $\eta^2 = 0.42$. Pairwise comparisons revealed that trust ratings significantly improved from before all autonomous journeys to after all autonomous journeys ($p < .001$), as well as from before all autonomous journeys to the unexpected stop ($p < .001$), and from before all autonomous journeys to the expected stop ($p < .001$). There was no significant difference in trust ratings between the unexpected and expected stop ($p = .10$).

For reliability ratings, a one-way repeated measures ANOVA model revealed a significant main effect for Journey, $F_{(1.44,\ 51.79)} = 25.55$, $p < .001$, $\eta^2 = 0.42$. Similarly to the trust ratings, pairwise comparisons revealed that reliability ratings significantly improved from before all autonomous journeys to after all autonomous journeys ($p < .001$), before all autonomous journeys to the unexpected stop ($p < .001$), and from before all autonomous journeys to the expected stop ($p < .001$). Again, there was no significant difference in reliability ratings between the unexpected and expected stop ($p = .10$). See Figure 5.4 and Table 5.2.

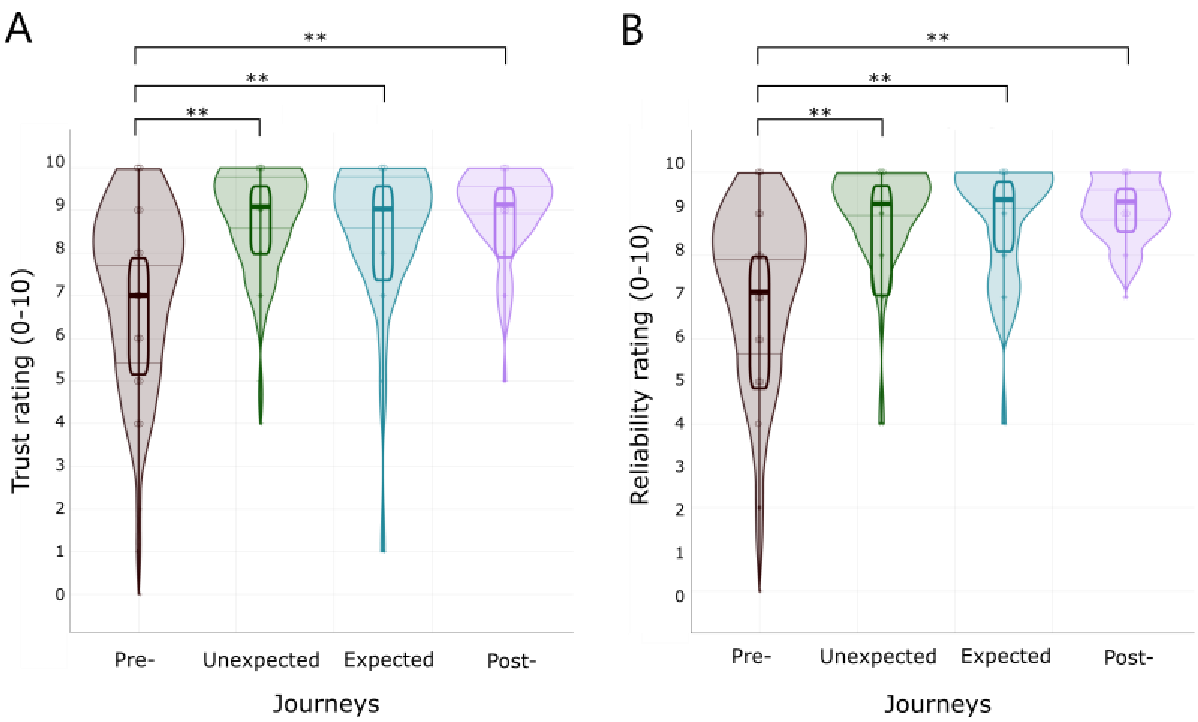



*Figure 5.4.* Violin plots representing trust and reliability ratings. **(A)** Trust ratings before all journeys, after the unexpected stop journey, after the expected stop journey, post-all journeys. **(B)** Reliability ratings before all journeys, after the unexpected stop journey, after the expected stop journey, post-all journeys. Key: ** $p < .001$.

*Table 5.2.* Mean (SD) of subjective ratings for trust and reliability before and after autonomous vehicle journeys.

| Subjective rating | Journeys | | | |
|---|---|---|---|---|
| | Pre- | Post- | After unexpected stop | After expected stop |
| Trust [0-10] | 7.12 (2.50) | 9.22 (1.13) | 9.16 (1.42) | 9.00 (1.78) |
| Reliability [0-10] | 7.19 (2.39) | 9.30 (0.88) | 9.35 (1.18) | 9.46 (1.02) |

### 5.4.2 Physiological arousal

*Heart rate:* The two-way repeated measures ANOVA was performed to understand the impact of an expected and unexpected stop on heart rate (*z*) activity during autonomous driving. The ANOVA revealed no significant main effects for Stop, $F_{(1, 29)} = 0.04$, $p = .85$, or TOI, $F_{(1, 29)} = 0.56$, $p = .46$. The interaction effect was also not significant, $F_{(1, 29)} = 0.06$, $p = .80$. Heart rate was similar between the period before an expected stop ($M = -.15$, $SD = 1.01$), and after an expected stop ($M = -.12$, $SD = .93$);

and between the period before an unexpected stop ($M$ = -.13, $SD$ = .95), and after an unexpected stop ($M$ = -.06, $SD$ = .92). As illustrated in Figure 5.5, heart rate rose during an unexpected stop, though this did not reach significance.

*Skin conductance level:* The two-way repeated measures ANOVA revealed no significant main effects for Stop, $F_{(1, 29)} = 0.17$, $p = .68$, or TOI, $F_{(1, 29)} = 0.37$, $p = .55$. However, the interaction effect between Stop and Time was significant, $F_{(1, 29)} = 0.98$, $p = 0.019$, $\eta_p^2 = 0.18$. Pairwise comparisons revealed that during the unexpected stop, skin conductance level ($z$) was greater during the 30 s after the stop ($M$ = 0.16, $SD$ = 0.83), compared to the previous 30 s ($M$ = -0.05, $SD$ = 0.72; $p$ = .04). There was no difference in skin conductance level ($z$) between the time before ($M$ = 0.06, $SD$ = 0.96) and after ($M$ = -0.09, $SD$ = 0.89) the expected stop. Figure 5.5 demonstrates that skin conductance increased after the vehicle stopped due to a hazardous event. This increase in arousal persisted for the 30 s following the stop.

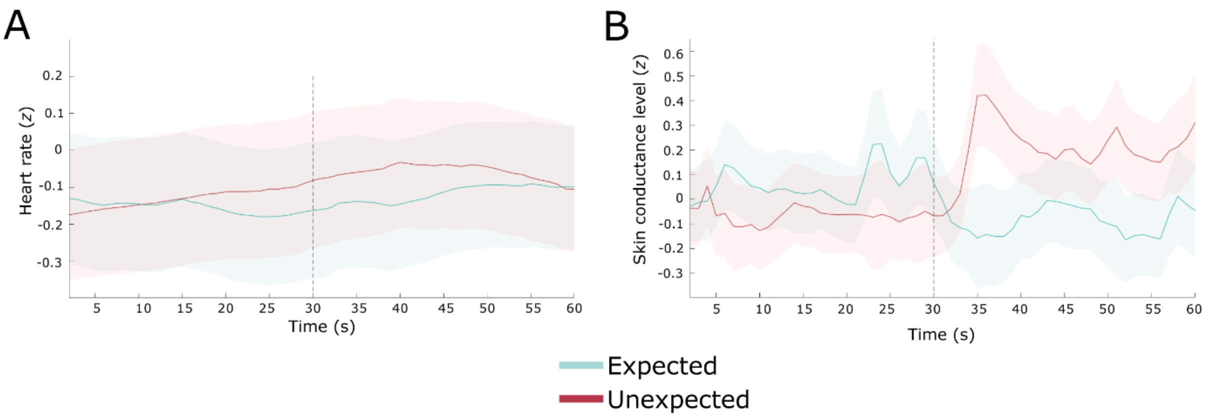



*Figure 5.5.* Heart rate and skin conductance level during an unexpected (red) and expected (blue) stop. Shaded areas represent the ± standard error of the mean difference. Grey dashed line represents the time point the vehicle stopped. **(A)** Heart rate ($z$). **(B)** Skin conductance level ($z$).

### 5.4.3 Eye tracking

*Fixation count:* To begin, a 2 (Stop) x 3 (AOI) repeated measures ANOVA was run to understand the impact of an unexpected stop on scanning behaviour of the visual environment. A significant main effect was found for AOI, $F_{(1.42, 35.55)} = 27.74$, $p < .001$, $\eta_p^2 = 0.53$; and the two-way interaction, $F_{(1.52, 38.10)} = 28.47$, $p < .001$, $\eta_p^2 = 0.53$. The main effect of Stop was not significant $F_{(1, 25)} = 0.44$, $p = .51$. *Post hoc* comparisons of the two-way interaction revealed a higher number of fixations in the central environment during an unexpected stop compared to an expected stop ($p < .001$); whereas fixation count was greater on the HMI during the expected stop compared to the unexpected stop ($p < .001$). Overall, in the unexpected stop, the number of fixations were higher on the HMI compared to the peripheral environment ($p < .001$), and the central environment compared to the peripheral environment ($p < .001$). Overall, during the expected stop, the number of fixations were higher on the HMI compared to the central environment ($p < .001$); and the HMI compared to the peripheral environment ($p < .001$). In combination, these results reveal that the number of fixations within the central environment were higher during an unexpected stop, whereas the number of fixations within the HMI were higher during an expected stop. See Table 5.3 for descriptive statistics.

Next, a 2 (Stop) x 3 (TOI) repeated measures ANOVA was run to understand whether visual scanning behaviour changed over time between stops. The main effects of Time, $F_{(1.56, 39.03)} = 0.88$, $p = .40$; and Stop, $F_{(1, 25)} = 0.44$, $p = .51$ were not significant, nor was the interaction effect, $F_{(1.21, 30.21)} = 1.45$, $p = .24$. See Table 5.4 for descriptive statistics, and Figure 5.6 and Figure 5.7 for an overview of the fixation count results.

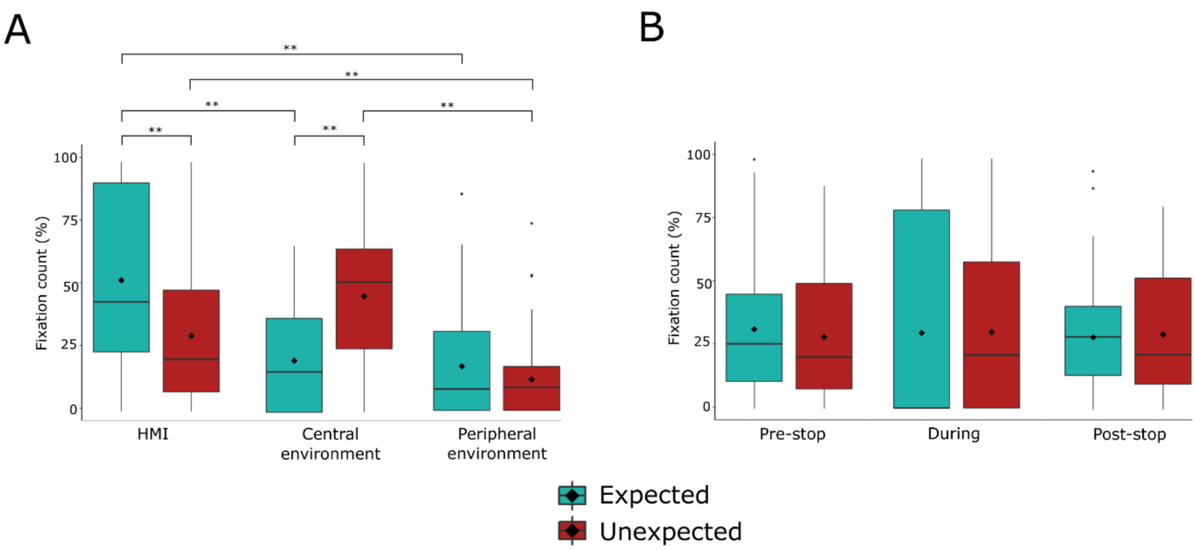



*Figure 5.6.* Fixation count (%) during unexpected and expected journeys. **(A)** Fixation count during unexpected and expected journeys over areas of interest (AOI). **(B)** Fixation count during unexpected and expected journeys over times of interest. Key: Bolded line represents the median value. Box represents the interquartile range. Vertical lines represent the lower/upper adjacent values. ◆ represents the mean value. ** $p < .001$; * $p < .05$.

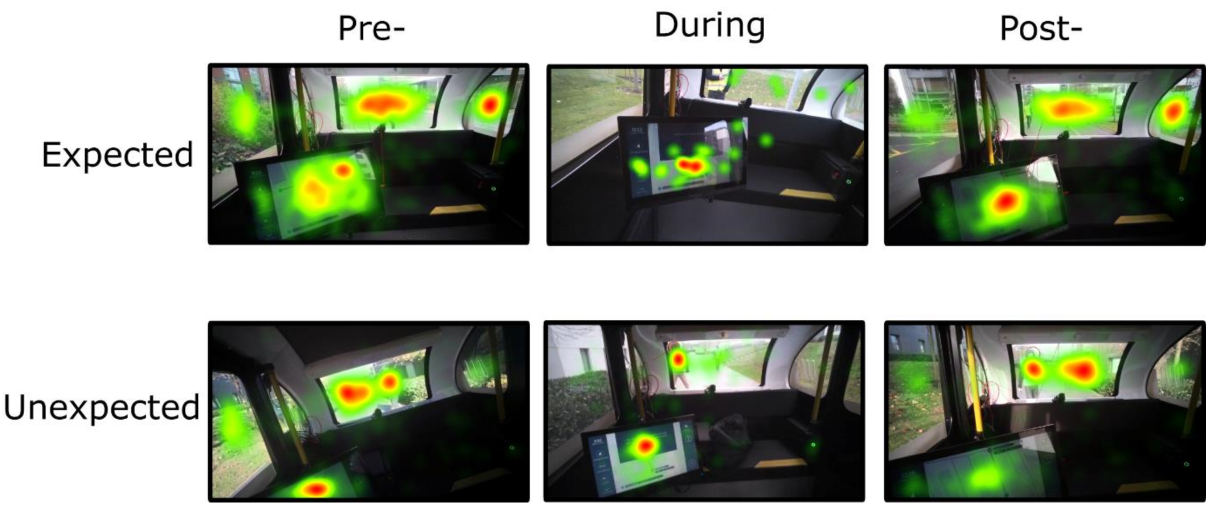



*Figure 5.7.* Total fixation count during an expected and unexpected journey. The heat map represents the summary of all gaze points in the visual environment over three time points of interest. Colours indicate the total gaze fixations (fixation count increases from green – yellow – orange – red).

*Fixation duration:* A 2 (Stop) x 3 (AOI) repeated measures ANOVA was run. The model revealed a significant main effect of AOI, $F_{(1.59, 39.72)} = 29.23$, $p < .001$, $\eta_p^2 = 0.54$; and a significant interaction effect, $F_{(1.58, 39.37)} = 23.27$, $p < .001$, $\eta_p^2 = 0.48$. The main effect of Stop was not significant, $F_{(1, 25)} = 0.44$, $p = .52$. *Post hoc* comparisons revealed that fixation duration on the HMI was longer during the expected stop compared to the unexpected stop ($p = .003$), but longer on the central environment during the unexpected stop compared to the expected stop ($p < .001$). Fixation duration on the peripheral environment was marginally greater during the expected compared to the unexpected stop ($p = .055$). Additionally, for the unexpected stop, fixation duration was shorter for the peripheral environment when compared to the HMI ($p <. 001$) and the central environment ($p < .001$). For the expected stop, fixation duration was longer on the HMI compared the peripheral environment ($p < .001$) and the central environment ($p < .001$). See Table 5.3 for descriptive statistics. Altogether, these results indicate that attentional demands were allocated towards the central environment during the unexpected stop, whereas attentional demands were allocated towards the automated systems interface during an expected stop.

Next, a 2 (Stop) x 3 (TOI) repeated measures ANOVA was run. The ANOVA yielded a significant main effect of Time, $F_{(1.63,\ 40.82)} = 6.52$, $p = .006$, $\eta_p^2 = .207$. The main effect of Stop, $F_{(1, 25)} = 0.44$, $p = .52$; and the interaction effect were not significant, $F_{(2, 50)} = 0.21$, $p = .813$. Fixation duration was greater during the stop ($M = 19.96$, $SD = 8.42$) compared to after the stop ($M = 16.48$, $SD = 7.72$), regardless of whether it was an expected or unexpected stop ($p$ = .01). See Table 5.4 for descriptive statistics, and Figure 5.8 and 5.9 for an overview of the fixation duration results.

In combination, the eye tracking results suggest that similar visual attention demands were expended but distinctly allocated during the unexpected and expected journey. Fixation duration was longer on the central environment during an unexpected stop, whereas fixation duration was longer on the interface during an expected stop, indicating distinct attentional resource allocation between the expected and unexpected stops.

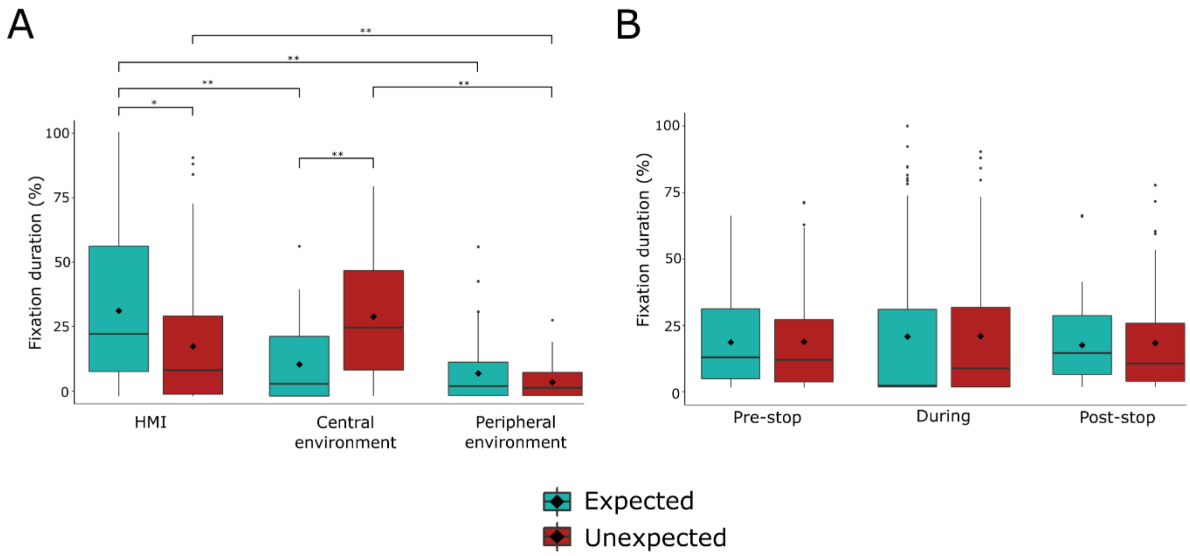



*Figure 5.8.* Fixation duration (%) during unexpected and expected journeys. **(A)** Fixation duration during unexpected and expected journeys over areas of interest (AOI). **(B)** Fixation duration during unexpected and expected journeys over times of interest (TOI). Key: Bolded line represents the median value. Box represents the interquartile range. Vertical lines represent the lower/upper adjacent values. ◆ represents the mean value. ** $p$ < .001; * $p$ < .05.

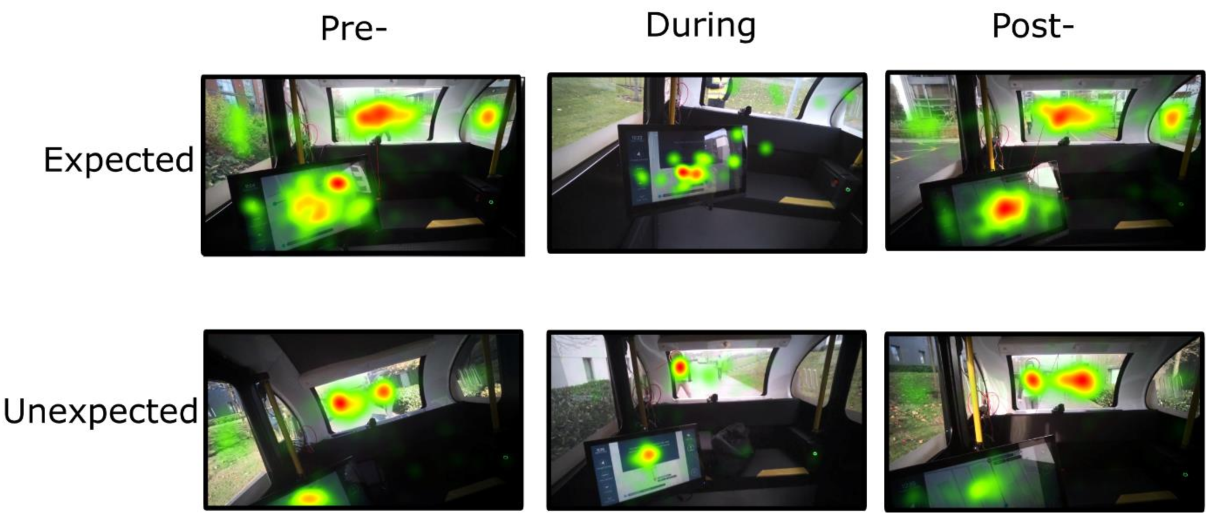



*Figure 5.9.* Total fixation duration during an expected and unexpected journey. The heat map represents the summary of all fixations in the visual environment over three time points of interest. Colours indicate the total gaze fixation duration (fixation duration increases from green – yellow – orange – red).

*Table 5.3.* Mean (SD) of fixation metrics count (%) and duration (%) across areas of interest the human-machine interface (HMI), central environment, and peripheral environment, during expected and unexpected stops.

| Fixation metric | HMI | | Central environment | | Peripheral environment | |
|---|---|---|---|---|---|---|
| | Expected | Unexpected | Expected | Unexpected | Expected | Unexpected |
| Fixation count (%) | 52.57 (30.07) | 30.32 (28.23) | 20.70 (20.02) | 46.50 (26.35) | 17.81 (20.85) | 12.52 (14.76) |
| Fixation duration (%) | 32.36 (27.75) | 18.86 (23.06) | 11.99 (13.86) | 30.03 (22.55) | 8.44 (11.54) | 5.07 (5.97) |

*Table 5.4.* Mean (SD) of fixation metrics count (%) and duration (%) across pre-, during, and post- expected and unexpected stops.

| Fixation metric | Pre- | | During | | Post- | |
|---|---|---|---|---|---|---|
| | Expected | Unexpected | Expected | Unexpected | Expected | Unexpected |
| Fixation count (%) | 31.94 (25.91) | 28.87 (25.68) | 30.17 (42.32) | 30.43 (31.74) | 28.97 (20.38) | 30.04 (25.03) |
| Fixation duration (%) | 17.36 (16.87) | 17.58 (17.79) | 19.34 (31.02) | 19.51 (25.3) | 16.08 (13.95) | 16.88 (19.27) |

## 5.5 Discussion

This study sought to understand the impact of an unexpected event during Level 5 autonomous driving on older adults' gaze behaviour, autonomic arousal, and self-report ratings of workload, perceived situation awareness and associated trust levels. To accomplish this, an experiment was designed where participants experienced what they thought were autonomous journeys that included two stops on separate journeys: one unexpected stop initiated by a 'hazard', and one expected stop initiated as part of the planned journey set up. Pre-journey versus post-journey ratings revealed that perceived situation awareness did not differ after periods of autonomous driving. Conversely, physical demand and effort factors decreased, while frustration increased. Trust and reliability ratings also increased from pre-journey values and remained high after each type of journey. Elevated electrodermal activity persisted after the unexpected stop. Gaze fixation metrics revealed several visual behaviour differences. Overall, participants searched the central environment, inclusive of the 'hazard', for longer during the unexpected stop, whereas during the expected stop, the human-machine interface (HMI) captured visual attention, as measured by greater fixation counts and longer fixation durations.

### 5.5.1 Autonomous driving impacted physical demand, effort, and frustration levels, but did not affect situation awareness

Subjective ratings of perceived situation awareness did not differ between pre- and post- journeys. This was unexpected as previous research has demonstrated modulated situation awareness during autonomous driving compared to manual driving and partially automated driving (De Winter et al., 2014). De Winter et al.'s (2014) meta-analysis revealed that situation awareness decreased if participants engaged with non-driving related tasks but improved if participants were notified or instructed through an HMI. However, in the present study, participants had no opportunity to engage in non-driving related tasks. In addition, due to time and cognitive demand constraints, the present study did not measure perceived situation awareness between conditions. Unlike the studies reviewed by De Winter et al. (2014), the rating technique was administered in a building separate to the autonomous vehicle environment. In addition, the studies reviewed also tended to measure situation awareness during tasks such as object detection. As De Winter et al. (2014) demonstrated that situation awareness is modulated by context-specific events, the present study unlikely captured participants' true situation awareness during autonomous journeys.

The physical demand factor of the NASA-TLX was lower after the autonomous journeys. This was expected, as autonomous driving does not recruit motor skills, such as moving the steering wheels or pedals. Findings also revealed lower effort values following autonomous journeys. The concept of effort can be considered as the amount of resources invested in the task. Although the driver did interact with the HMI to set up journeys, effort was not relevant for maintaining successful journeys. As mental workload did not differ between conditions, this provides further evidence that workload is a multidimensional construct. Whereas mental workload can be considered the amount of cognitive demand required; effort represents how much one must work to achieve something. Being in a dynamic visual environment may have engaged participants' cognitive mechanisms, however, did not experience significant demand on their cognitive processes.

Interestingly, frustration significantly increased following autonomous journeys. Although not greatly researched, this is similar to previous studies that have found that warnings during automation increase frustration (Kassner et al., 2011), and when automated systems cannot be overridden increase frustration (Comte, 2000). Similar to situation awareness, the present study only measured workload factors before and after all journeys, so it is not clear what aspects of the autonomous journeys may have caused an increase in frustration levels.

### 5.5.2 Trust and reliability ratings were high after each journey, regardless of the stop

Subjective ratings of trust and reliability remained high between the unexpected stop with a hazardous event and the expected stop without a hazardous event. Although participants had no previous experience with an autonomous vehicle, trust ratings averaged 9 out of 10 throughout the journeys. Trust has shown to be an important concept that plays a leading role in the willingness of humans to rely on automated systems during critical events (Hoff & Bashir, 2015). Overreliance on an automated system can be subsequently harmful to a user if they accept the automated system's recommendations and actions inappropriately, when system errors do occur (Parasuraman & Riley, 1997). In the present study, several reasons may have encouraged high trust levels despite contextual differences between journeys. Choi & Ji (2015) suggest that technical competence is an important factor that encourages high levels of trust in autonomous vehicles. Technical competence refers to user perception on the performance of the vehicle. Potentially, the vehicle behaving appropriately to the hazardous event (e.g. braking and notifying the participant) may have encouraged users to overtrust the vehicle's behaviour once the journey had successfully completed. In addition, a lack of system information, including a lack of direction from the HMI, may have encouraged users to overtrust the vehicle's behaviour once the journey successfully completed with no system errors. As older adults are more likely to rely on automated systems (McBride et al., 2011) and are more likely to lack understanding of advanced technology (Mann et al., 2007), the limited system transparency could have increased trust while the vehicle performed as expected. Studies have found that the right amount of system information is paramount to achieve the optimum level of trust between human and machine (Choi & Ji, 2015). Overall, these results indicate that facilitating appropriate trust levels in older adults for autonomous vehicles is important to improve safe human-vehicle interaction.

### 5.5.3 Skin conductance levels increased after an unexpected stop

The results revealed that skin conductance levels increased when the vehicle was presented with an unexpected event – the vehicle coming to a stop due to a hazard on the road. This increase persisted when the vehicle restarted. However, skin conductance levels did not differ between the unexpected and expected stop. Driving studies have indicated that high EDA levels are modulated by workload (e.g. Mehler et al., 2012), stress (e.g. Affanni et al., 2018), lower trust in automation (e.g. Morris et al., 2017), and anxiety (e.g. Barnard & Chapman, 2018; see Lohani et al., 2019 for a review). For example, Mehler et al. (2012) found that skin conductance levels were higher during dual-task driving compared to single-task driving, demonstrating an increase in cognitive load associated with elevated skin conductance levels. The authors also found that during autonomous driving, skin conductance levels increased when compared to conventional manual driving. This work emphasises that EDA is sensitive to a variety of psychological phenomena, and therefore caution should be exercised when interpreting changes in EDA in less controlled settings (Dawson et al., 2007). It is therefore difficult to infer specifically why skin conductance levels rose, other than reflecting an overall increase in sympathetic arousal. Trust ratings were high after all journeys, implying trust levels did not modulate sympathetic arousal. However, whether the increase in sympathetic arousal was associated with time-varying trust factors, workload, or anxiety, cannot be indicated with physiological measures alone.

There was not a statistically significant difference in heart rate, although Figure 5.5 shows elevated heart rate, for the unexpected stop compared to the expected stop. As skin conductance is regulated by the sympathetic nervous system, and heart rate is modulated by both the activation and suppression of sympathetic and parasympathetic branches of the autonomic nervous system respectively (Thayer et al., 2010), the results suggest that the vehicle stopping in response to a hazardous event reflects a mild sympathetic dominance. Ruscio et al. (2017) measured physiological responses to takeover requests following various warnings. Heart rate decreased relative to manual driving following reliable warnings, misleading warnings, and no warnings. Skin conductance response amplitude increased during a misleading warning and no warning. They also found that respiratory sinus arrhythmia, an index of parasympathetic activity, increased from manual driving to an unexpected takeover with no warning. Their results reveal an imbalance between the parasympathetic and sympathetic branches during takeovers preceded by a misleading warning or no warning. The authors suggest that this discrepancy may reduce attentional capacity, resulting in cognitive overload. Although the present study did not measure specific or nonspecific response amplitude changes, but rather changes in skin conductance level, the results are somewhat compatible as the present study found a similar effect of increased skin conductance level following vehicle cessation without any warning (unexpected stop). In addition, the effect size was similar albeit slightly smaller than Ruscio et al. (2017) indicating a moderate effect ($\eta_p^2$ = .176 versus $\eta_p^2$ = .196). However, the results are difficult to directly compare to Ruscio et al.'s (2017) findings directly as the present study did not measure physiological responses during manual driving or initiate a takeover request, nor did it separate parasympathetic activity from sympathetic activity; therefore, it is not clear whether a reduction in attentional capacity was associated with a marked increase in sympathetic activation as measured by an increase in skin conductance level.

### 5.5.4 During an unexpected stop, visual attention was directed towards the hazard compared to the peripheral environment and human-machine interface during an expected stop

The distribution and duration of fixations measured with eye tracking outline the differences between the two types of journeys: while in the unexpected stop journey participants' fixation duration was longer and fixation count was higher towards the central environment (containing the 'hazard'), in the expected stop visual attention was directed toward the HMI. In addition, the results revealed that participants' visual allocation was more diverse before the expected stop, as indicated by the greater distribution of fixations across the scene, when compared to the unexpected stop. Visual behaviours between stops were similar over time, indicating similar demands on visual attention. The effects were highly significant and associated with strong effect sizes.

General visual scanning behaviour can be understood by fixation counts, as more shifts within a scene are associated with a greater frequency of fixations. Fixation duration can provide further insight by indicating visual attention demands. Basic visual processing research has demonstrated that fixation duration increases with visual scene complexity (e.g. Pomplun et al., 2013) and cognitive load (e.g. Rayner, 1998), and is linked to uncertainty (e.g. Brunyé & Gardony, 2017). As such, the patterns of results imply that during an unexpected stop, visual attention was directed toward the central environment containing the 'hazard,' with the participants searching for information, rather than focusing on other aspects of the scene. This may be due to increased demand placed upon the vision and attention systems by the visual scene complexity and uncertainty. The results are similar to research studying ocular behaviour during manual driving and hazardous situations. The variance of fixations decreased when presented with a critical situation (Chapman & Underwood, 1998). In contrast, fixation duration increased coming up to, and during, a critical situation (e.g. Chapman & Underwood, 1998; Underwood et al., 2005). In addition, a negative relationship between task demands during driving and visual scanning behaviour has been demonstrated, i.e., higher task demands reduced the dispersion of visual scanning (e.g. Recarte & Nunes, 2003; Savage et al., 2013). Moreover, Guo et al. (2019) found that fixation frequency and duration increased during accident scenes reflecting increased anxiety. Searching for information related to the unexpected event might be explained by increased anxiety (Guo et al., 2019): the narrowing of visual attention, such as focusing on a hazard, is a common feature of increased arousal and stress (Chajut & Algom, 2003; Gable & Harmon-Jones, 2010). Moreover, the physiological results show increased sympathetic arousal following the unexpected stop.

Although the results are supported by the above-mentioned driving studies, the present study differs significantly as the participant was not an active manual driver in control of the vehicle. As ocular behaviour and motor execution are intrinsically linked both spatially and temporally, active drivers successfully fixate directly at the objects being interacted with or ones that precede the action. Despite these differences, the results are in agreeance with Strauch et al. (2019) who investigated eye gaze of passengers during real-world autonomous driving. They found a greater frequency of fixations on safety-relevant areas of interest when joining a highway during an autonomous journey when compared to manual driving and the rest of the route. In accordance with these studies, visual scanning behaviour was affected by safety-critical situations regardless of active involvement in the driving task. In addition, Strauch et al. (2019) participants' mean age was 23 years, and so the present results extend this earlier research suggesting that older adults display similar ocular behaviours to younger adults during safety-critical situations. The present study does differ as passengers interacted with an HMI throughout the journeys. Despite this, the results found that attentional focus narrowed toward the 'hazard' before, during, and after the unexpected event, which was also accompanied by an increase in skin conductance reflecting increased sympathetic nervous system arousal following the vehicle response. It should be noted that the narrowing of visual attention is most likely associated with object-based attention, as participants focused on an object defined in a spatial frame, i.e. the hazardous person, rather than exclusively spatial or feature-based attention.

### 5.5.5 Limitations

Although the current study attempts to produce increased ecological validity compared to the laboratory studies, safety restrictions were put in place including the speed of the vehicle, the safety driver, and the marshals surrounding the vehicle. On average, the vehicle speed was approximately 3 to 5 mph. The speed of a vehicle has been known to correlate with self-report workload measures i.e. the greater the speed, the greater the self-reported workload (Fuller, 2005). However, research has found that this depends on the situation complexity. Low complexity environments including motorways at faster speeds, or high complexity situations including town centres at lower speeds, may modulate load in a similar manner (Paxion et al., 2014). In this study, the vehicle drove around a pedestrianised area, where the maximum speed limit was 10 mph. The vehicle shared the lane with many pedestrians, cyclists, and obstacles such as bollards. Therefore, driving at a greater speed would not have been possible nor realistic, even during manual driving.

A potential limitation was the use of the Empatica E4 for assessing autonomic arousal. Gruden et al. (2019) recently found that manual driving-related movement artefacts impacted heart rate variability and skin conductance level measurements taken by the E4. Reasonable accuracy and reliability have been reported for this device providing wrist movements are low (Ragot et al., 2017), which was the case during the present study, as Level 5 driving does not require behaviours such as changing gears. Nevertheless, Appendix 5.4 demonstrates additional analyses that confirm accelerometer values did not differ between conditions. Therefore, the differences in skin conductance level cannot be attributed to differences in motor activity. The issue of movement artefacts is discussed further in Chapter 7.3.3.

Finally, the results imply that the unexpected event placed significant demands on attentional resources. However, eye tracking is an indirect measure of attention, and as the study mimicked Level 5 autonomous journeys, no direct performance measure could be derived to support this view. Yet, all participants were introduced to a “Safe stop” button on the HMI, which could be pressed at any time if they wanted the vehicle to stop. None of the participants activated the safe stop. They could also have accessed the “Vehicle health” icon, which would have presented them with information about the overall health of the vehicle. None of the participants accessed this icon during the times of interest. Taken together, these findings suggest that the unexpected stop was not perceived as particularly dangerous as either subjective ratings, or all physiological indices reflected an extreme response that may be associated with more imminent or extreme danger. Further investigations using different types of unexpected events are needed to be able to characterise functional states to specific safety-critical scenarios.

### 5.5.6 Conclusion

Taken together, these results have several critical implications for the safe implementation of Level 5 autonomous vehicles for older adults. The results reveal possible narrowing of visual attention and heightened arousal during an unexpected event as demonstrated by increased sympathetic arousal and a smaller distribution of fixations, coupled with an increase in fixations toward the unexpected event. In combination with consistently high trust ratings, these results suggest that the passive process of automated driving may restrict the focus of visual attention and heighten adverse responses. This study also demonstrates that the physiological indices examined can be useful and practical measures for evaluating passive drivers' functional state during real-world semi-autonomous and fully autonomous driving. As such, a driver state monitoring (DSM) system that includes physiological indices might be able to detect these behaviours and make an informed decision on vehicle behaviour and adapt HMI notifications accordingly. For example, when attention is negatively impacted in a safety driver responsible for monitoring of automation errors and hazards during the development of a fully autonomous vehicle, or passive drivers required to takeover in semi-autonomous vehicles. In addition, the potential for negative experiences during autonomous driving, coupled with human limitations in sustained monitoring during low and high arousal situations, suggests that a DSM system may be a necessary adjunct to fully autonomous vehicles in supporting the health and wellbeing of potentially vulnerable people in unexpected situations. Therefore, the final study described in Chapter 6 aimed to investigate the impact of a DSM notification on visual attention and autonomic arousal during autonomous driving, as well as exploring potential barriers towards acceptance.

# 6.0 The Impact of a Driver State Monitoring Notification on Functional State During Simulated Autonomous Driving

## 6.1 Overview

Chapter 5 revealed potential negative experiences during autonomous driving coupled with human limitations in sustained monitoring during high arousal situations, suggesting a driver state monitoring system may be a necessary adjunct to fully autonomous vehicles to improve safety and wellbeing. Therefore, Chapter 6 employed a study designed to understand the impact of a biological driver state monitoring notification on visual attention and autonomic arousal during autonomous driving, as well as exploring potential barriers towards acceptance. Participants undertook several simulated autonomous journeys. Autonomic arousal and visual attention via eye tracking fixation metrics were compared during two distinct notifications that preceded the vehicle slowing down. The first notification displayed biofeedback changes in physiological state; the second notification provided speed limit changes. Self-report changes in workload, situation awareness, trust, and reliability were also collected, as well as acceptability ratings.

## 6.2 Introduction

Chapters 2, 3 and 4 demonstrated possible attentional deficits as indicated by neurophysiological indices, autonomic arousal, and task performance during multitasking and task switching, while Chapter 5 revealed modulations in sympathetic arousal and ocular behaviour during an unexpected event during automated driving. As a result, allocating the appropriate cognitive resources to process information relevant towards automated driving could be impacted by the current driving environment (e.g. an unexpected event), previous attentional load (e.g. engaging in an unrelated task), and task switching (i.e. responding to an automated vehicle's unexpected takeover request or vehicle notifications), resulting in significant implications for the safe development of autonomous vehicles and the health and wellbeing of passengers in autonomous vehicles. As such, it is argued that an advanced driver assistance system could determine the extent to which the safety driver in a fully autonomous vehicle or a passive driver in a semi-autonomous vehicle is suitable for the current driving scenario by monitoring the driver's functional state as indexed by physiological indices (e.g. Rauch et al., 2009).

### 6.2.1 Potential positive effects of driver state monitoring

Driver state monitoring (DSM) is not a new or unique topic. Manual driving has benefited from DSM systems to detect fatigue and inattention to improve vehicle safety (Melnicuk et al., 2016). Situations such as night-time driving (Phipps-Nelson et al., 2011), prolonged driving (Finkelman, 1994), and extreme temperatures (Xianglong et al., 2018) can induce fatigue; whereas mobile phones (Strayer & Drews, 2007), in-vehicle systems (Arexis et al., 2017), and eating (Tay & Knowles, 2004) can induce inattention. Researchers have combined a hybrid of measures to attempt to accurately detect deficits in functional state. Subjective report measures (e.g. self-rating of current fatigue level), driver biological measures (e.g. muscle activity), driver physical measures (e.g. blink frequency), and driver performance measures (e.g. lateral position) have been used to measure deficits (Dong et al., 2011). For example, Liang et al. (2007) detected driver distraction with 81% accuracy utilising a combination of eye metrics and vehicle behaviours such as eye fixation, smooth pursuit, steering-wheel angle, and lane position. Jacobé et al. (2019) recently combined several physiological, behavioural, and vehicle measurements, and found that a trained artificial neural network detected fatigue performance decrements within an accuracy of five minutes. Including more information, such as context (e.g. traffic, type of road), improved accuracy further.

The implementation of a DSM system has many potential benefits for improving vehicle safety and passenger wellbeing during semi-autonomous and fully autonomous driving. Although the overall aim of the system is the same as manual driving, to improve vehicle safety it will need to detect additional adverse behaviours due to the passive role of the driver. During the development of Level 4 and 5 vehicles, the information about a safety driver's state could be used to modify in-vehicle information. For example, the system could present the optimum amount of feedback or information in regard to their current cognitive load (i.e. underload or overload). For example, it could choose between auditory or visual feedback depending on what the driver is doing or how they are feeling. Over time, it could predict lapses in attention and prevent them from occurring. This can be also considered an important role for a DSM system in semi-autonomous driving, when at times a passive driver must takeover from the automated driving system. In addition, the system could improve the wellbeing of a passenger in an autonomous vehicle. The system could track the influence of a driving situation on a passenger's state and estimate the severity on their wellbeing. In turn, the vehicle could adapt its behaviour to improve a passenger's wellbeing e.g. leaving a greater headway between the vehicle in front.

### 6.2.2 Potential negative effects of driver state monitoring

Despite the benefits and willingness of the research community to adopt DSM systems, an effective system must have the desired outcome on functional state. The system must be appropriately designed and adequately implemented for users to accept it. In addition, a DSM system may have the reverse anticipated effect on driver state. Driving research has demonstrated that non-driving related notifications have a negative impact on driving performance. For example, Giang et al. (2015) found that notifications received on a smartwatch resulted in a delayed braking response during a critical event. Even the mere presence of receiving a notification has been shown to negatively impact driver performance (Stothart et al., 2015). As such, notifications that demand users' attention at inconvenient moments are likely to have adverse effects and divide attention, resulting in disruption rather than proving beneficial.

However, a DSM system will evoke a notification related to the task, rather than non-driving related. Occupational research investigating human-machine interaction can therefore provide further insight into the effects of task-related notifications. Giraudet et al. (2015) monitored P3b response to an auditory oddball task during an air traffic control task. Two types of visual notifications were displayed: one salient notification which highlighted the text, and one non-salient notification which did not highlight the text. Participants were more accurate to detect the salient notification, and demonstrated a greater P3b when compared to the less salient notification. Therefore, a salient visual notification resulted in a lower depletion of attentional resources required for auditory processing. As auditory alarms are sometimes not perceived, particularly if visual processing load is high (also known as inattentional deafness, Macdonald & Lavie, 2011), this has important implications for DSM systems. If visual load is high, users may not attend to an auditory notification therefore any feedback must be designed appropriately to ensure it catches the passive driver’s attention under different demands of visual load.

The literature has recently provided further insight into the impact of notifications during periods of automated driving. Ulahannan et al.’s (2020) study investigated ocular attention towards a human-machine interface during partially automated driving. Over five days, participants were more likely to monitor the roadway rather than direct their gaze towards an in-vehicle display. An important aspect of monitoring the automation was to engage with the system, as crucial information such as the technical competence of the vehicle and features such as hazard detection were displayed on the vehicle’s system. Therefore, the authors suggest that is it vital that the system can engage attention appropriately and encourage automation monitoring behaviours. However, Ulahannan et al. (2020) only investigated steady-state driving and did not vary notifications over time. Therefore, it is not clear whether specific notifications designed to engage attention could influence eye gaze. Yet, a recent study can provide some insight as the authors did vary notifications over simulated autonomous journeys. Eimontaite et al. (2020) demonstrated reduced heart rate and skin conductance levels during a journey containing audio and visual notifications when compared to a journey without audio notifications. In addition, subjective workload was lower during audio and visual notifications compared to visual only notifications. The authors suggest that multimodal notifications reduced arousal and workload mechanisms as participants may have felt they were kept ‘in-the-loop’ with audio notifications, diminishing any negative feelings of stress or anxiety. Despite lower arousal, it is unclear whether these effects negatively impacted attentional resources, as previous studies have demonstrated worse performance during presentation of multimodal notifications (e.g. Stothart et al., 2015).

Feedback of physiological signals could induce feelings of negative affect and evoke state anxiety in drivers. Not only may this have a negative effect on a person’s wellbeing, but anxiety may also have a negative impact on attentional networks the DSM system is attempting to engage. Pacheco-Unguetti et al. (2010) found that state anxiety impacted bottom-up processing of alerting and orienting, increasing reaction times. In addition, anxiety has been found to increase activity in the superior temporal sulcus (Bishop et al., 2007), an area also critical for attention modulation (Bogadhi et al., 2019). Together, these results suggest that it is important a DSM system conveys the right amount of information to successfully modulate attentional resources without negatively impact wellbeing.

### 6.2.3 Experiment rationale

It is important to understand how notification of changes in driver biological state might impact attention and wellbeing during autonomous driving. In addition, for DSM to be successful, the system must be well-accepted by the intended users. Therefore, the aim of the study was to investigate whether a DSM message could have detrimental effects on ocular attention and autonomic arousal, as well as user acceptability, during simulated Level 5 driving. To investigate this, participants experienced two types of journeys that presented a distinct notification. The first notification displayed biofeedback changes in physiological state; the second notification provided speed limit changes. Real-time signals of physiological arousal and eye gaze, as well as retrospective trust, reliability, and situation awareness ratings were compared. Users’ self-reported perception of the

DSM system were also collected. Understanding user perception was exploratory in nature and was carried out as an initial step to inform further research.

Previous research has indicated that salient task-related notifications increase attentional resource allocation (Giraudet et al., 2015). Therefore, it was expected that the DSM notification would direct visual attention towards the roadway, resulting in longer fixation duration and a greater frequency of fixation, when compared to a speed limit change notification. At notification onset, first fixation metrics towards the notification should be greater for the DSM notification compared to the speed limit change notification. It was expected that physiological arousal would increase following notification onset, reflecting an alerting mechanism, which would recover quickly. As salient notifications increase the feeling of being in control and reduce autonomic arousal (e.g. Eimontaite et al., 2020), it was expected that electrodermal activity and heart rate would be lower, and heart rate variability would be greater, following a DSM notification when compared to a speed limit change notification. Subjective measures of trust, reliability, situation awareness and workload were collected to further investigate the relationship between DSM notifications and autonomic arousal and attention.

## 6.3 Method

### 6.3.1 Participants

Thirty-three young adults originally participated in this study. Two participants were excluded from all subsequent analyses due to the simulator experiencing technical errors, leaving 31 participants (nineteen females, twelve males, mean age ± SD = 20.39 ± 2.25 years, range 18-25 years). 28 participants physiological and eye tracking data were subsequently analysed (eighteen females, ten males, mean age ± SD = 20.07 ± 1.53 years, range 18 – 25 years). This was due to recording errors with the eye tracker (two participants), and the event marker not accurately capturing the onset of conditions (one participant).

Individuals with severe motion sickness, severe health conditions (i.e. epilepsy, neurological impairments) and uncorrected vision or hearing were excluded. Similarly, those who did not have a high level of spoken and written English competency were not included, as the notifications were presented in English. Due to eye tracking measurements, individuals with eye movement abnormalities or who had undergone eye surgery could not be included in the study. Ethical approval was obtained by the Faculty of Health and Applied Sciences University of the West of England Research Ethics Committee (HAS.16.10.026). All participants gave written informed consent and were fully debriefed at the end of the study.

### 6.3.2 Apparatus

#### *6.3.2.1 Driving simulator*

The semi-autonomous simulation environment consisted of a Lutz Pathfinder Pod and three large forward projector screens with a display resolution of 1280 x 1024 providing a 210° horizontal forward field of view. The driving simulation was generated by the SCANeR II® software (OKTAL Sydac, France) and consisted of a two-lane road passing through a neighbourhood with limited traffic passing in the opposite direction. The scenarios included avatars walking along the pavements, pedestrian crossings, other vehicles, and buildings such as houses and cafes. Audio of simulated engine, road, and traffic sounds were presented. See Figure 6.1 for a depiction of the driving environment.

The static simulator consisted of a Lutz Pathfinder pod which has two forward facing seats with two rear-hinged doors. The vehicle does not include a steering wheel, pedals, or a gear stick. The participant was told to interact with the vehicle using the in-vehicle human-machine interface (see Figure 6.2).

*Figure 6.1.* The simulated driving environment. The driving simulation was a rural environment consisting of a two-lane road passing through a neighbourhood with limited traffic passing in the opposite direction.

*6.3.2.2 Human-machine interface (HMI)*

The human-machine interface (HMI) was presented on a HannsG HT161HNB 15.6" Multi Touch Screen connected to a Kodlix GN41 Mini PC (Windows 10, Intel Celeron processor, 8/4GB RAM, 64GB). The design of the HMI was similar to Chapter 5 (see Chapter 5.3.2.2), however the design had been updated to include additional user options such as ability to turn on and off voice notifications and visual display settings (i.e. font size, colour, and layout). See Figure 6.2 for an example of the HMI layout.

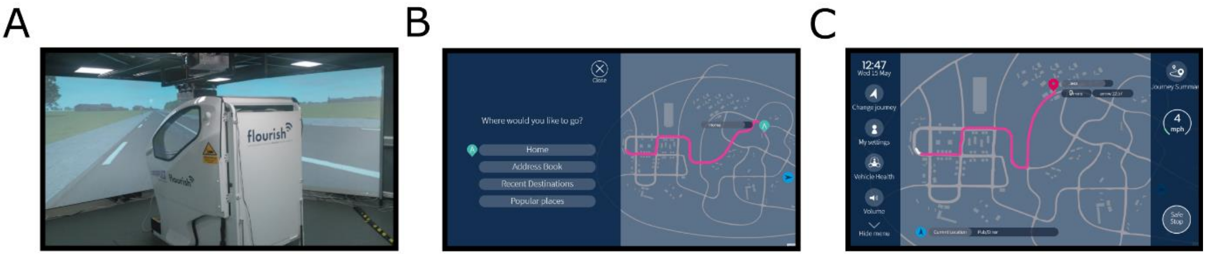



*Figure 6.2.* The driving simulator and human-machine interface (HMI). **(A)** The exterior of the Lutz Pod in front of three large forward projector screens. **(B)** The HMI screen during journey set up. **(C)** The HMI default screen during the journeys.

## 6.3.3 Journeys

Participants partook in a total of eight randomised journeys. Before each journey, participants were provided with a scenario that specified where they were, where their destination was, and if they were required to stop anywhere in between. The eight journeys always consisted of a journey with a speed limit change message, and a journey with a driver state monitoring (DSM) message. These notifications were pre-programmed and occurred at the same time during each journey. Once the messages appeared, the vehicle reduced its speed from 30 mph to 10 mph. The participant was not provided with any opportunity to turn the notifications off, nor could they increase the speed of the vehicle. After approximately 20 seconds (depending on traffic and traffic lights), the notification disappeared, and the vehicle returned to the normal speed limit. The notifications always occurred in the final four journeys and were randomised between participants to prevent order effects. Each journey was of approximately 10 minutes duration, and so participants experienced 80 minutes of the driving simulator in total.

Participants experienced two journeys that manipulated notification type. During one journey, a speed limit change message was presented on the HMI, *"Approaching the town centre. Speed limit change".* During another journey, a driver state monitoring (DSM) message was presented, *"Your physiology suggests that your reaction times are slowed. The vehicle will slow down".* Before the DSM journey, participants were told that their physiological state was being monitored to assist vehicle safety, via the Empatica E4 wristband, and reminded that they could press the 'Safe stop' button at any time. In actuality, participants' real-time physiological state was not being monitored and the DSM message was presented at the same location during the journey for each participant.

### 6.3.4 Protocol

On arrival, participants were provided an information sheet, filled in a consent form and a demographic questionnaire. Participants received a journey familiarisation instruction sheet that contained an overview of the journey tasks. After filling in pre-journey questionnaires, participants were seated in the simulator and experienced virtual journeys while interacting with the HMI. To begin, participants were told to keep a look out for any obstacles and press the 'Safe stop' button on the HMI if they needed to. Participants were required to set up each journey, entering a destination and a proposed stop. After each journey, participants completed after-journey measures and were provided with a scenario description for the next journey. After all journeys were completed, participants completed post-journey questionnaires. The testing session lasted for approximately 170 minutes depending on inter-subject individual differences. See Figure 6.3 for a schematic of the experimental procedure.

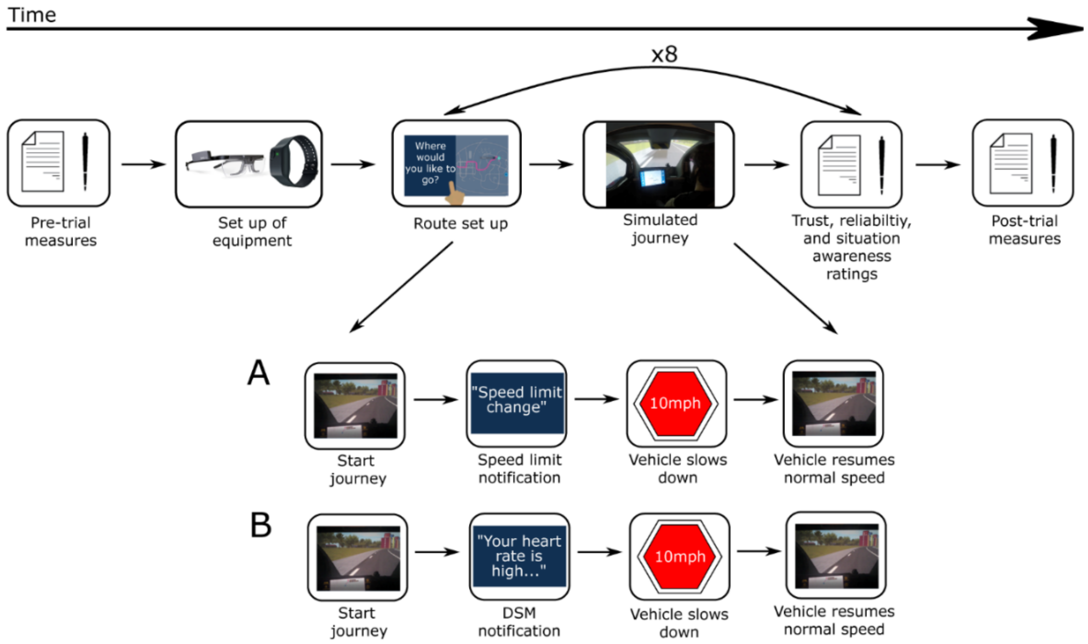



*Figure 6.3.* Experimental procedure. Journeys always consisted of a speed limit change journey and a driver state monitoring (DSM) journey. **(A)** During the speed limit change notification, a notification was presented on the HMI, *"Approaching the town centre. Speed limit change".* The vehicle then slowed down for approximately 20 seconds before returning to the normal speed. **(B)** During the DSM journey, a notification was presented, *"Your physiology suggests that your reaction times are slowed. The vehicle will slow down"*. The vehicle slowed down for approximately 20 seconds before returning to the normal speed limit.

### 6.3.5 Measures and pre-processing

This study involved a similar combination of measures which were administered during the applied Pod study described in Chapter 5 (see Section 5.3.5.1 for an overview of self-report measures). Pre-processing steps which differ to Section 5.3.6.1 are described below.

#### *6.3.5.1 Self-report measures*

*The NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988):* The NASA-TLX was administered to measure workload factors: mental demand, physical demand, temporal demand, effort, performance, and frustration. Participants filled in the questionnaire before and after all simulator journeys. Similar to Chapter 5, the NASA-TLX factors were averaged independently to derive the raw-TLX score for each subscale (mental, physical, temporal, effort, performance, frustration).

*Situation Awareness Rating Technique (SART; Taylor & Selcon, 1990):* The SART was administered to measure perceived situation awareness. In addition to participants filling in the questionnaire before and after all simulator journeys, they were asked to fill in the questionnaire after each journey. Similar to Chapter 5, an overall SART score was derived using the following formula: summed understanding divided by (summed attentional demand – summed attentional supply).

*Trust and reliability ratings:* Single-item trust and reliability ratings were collected. Similar to Chapter 5, participants completed these ratings before and after all simulator journeys, as well as after each journey. In this study, participants provided a written, rather than verbal, response.

*6.3.5.2 Physiological arousal*

Physiological measures of heart rate (beats per minute; BPM) and electrodermal activity (skin conductance level; µS) were collected using an Empatica E4 wristband (Empatica Inc., Cambridge, MA, USA and Milan, Italy) to measure levels of autonomic arousal. Please see Chapter 5 subsection 5.3.6.2 for further details about this measure, and 5.3.7.2 for the pre-processing steps. Similar to Chapter 5, missing interbeat interval samples from the Empatica meant it was not possible to calculate accurate heart rate variability measures RMSSD, SDSD, and pNN50, and so these metrics were not further analysed. Data were averaged within two times of interest: 30 seconds before the notification and 30 seconds after the notification.

*6.3.5.3 Eye tracking*

Tobii Pro Glasses 2, an eye tracking device, was used to collect fixation metrics (Tobii Glasses Eye Tracker, Tobii Technology, Stockholm, Sweden). Please see Section 5.3.5.3 for further details about this measure. Similarly to Chapter 5, individuals wearing glasses were asked to remove them, and a suitable prescription lens was attached to the glasses (ranging from –5.0 dpt to +3.0 dpt). Once the participant was wearing the head unit, the manufacturer's calibration procedure was followed which typically took less than 30 seconds and consisted of participants fixating on a central target. Eye tracking glasses captured a mean of 78% (*SD* = 11%) of gaze samples overall. Times of interest (TOI) were defined as 30 seconds before the notification, during the notification, and 30 seconds after the notification. Areas of Interests (AOIs) were defined on each mapped image for each TOI (see Figure 6.4). Two AOIs were created representing the HMI and the driving environment. Section 5.3.6.3 provides specific details about the pre-processing steps.

Time to first fixation, duration of first fixation, total fixation duration, and total fixation count metrics were exported. Because the notification period varied across participants, it was necessary to calculate fixation count and fixation duration proportions based on the total numbers of fixations and fixation durations. Fixation duration was defined as the amount of time spent looking at each AOI divided by the total duration of fixations. Fixation count was defined as the number of fixations towards each AOI divided by the total number of fixations.

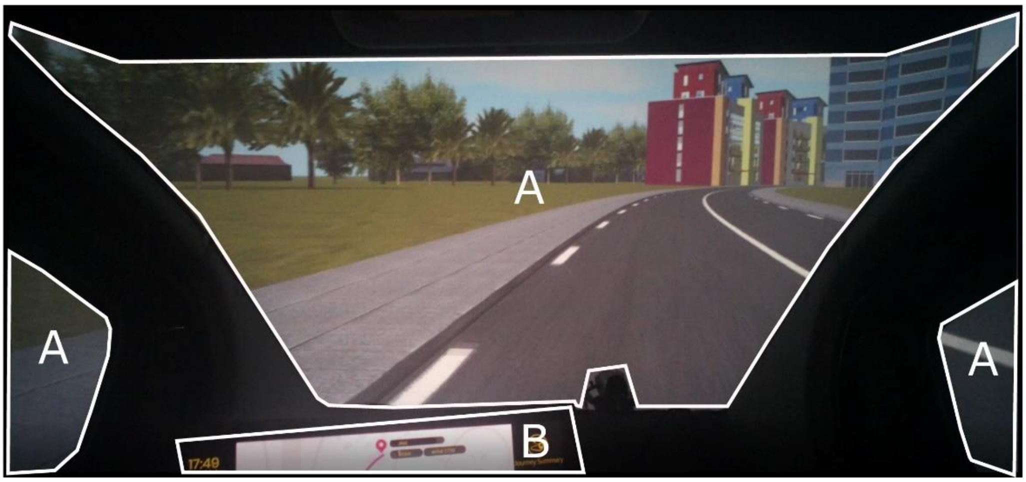



*Figure 6.4.* Areas of Interest (AOI) for eye tracking analysis. **(A)** Driving environment. **(B)** Human-machine interface.

### *6.3.5.4 Subjective experience of the driver state monitoring system*

During the post-trial questionnaire phase, participants filled in several questions regarding the DSM system. Perceived usefulness of the system was measured using the usefulness dimension of the Usefulness, Satisfaction, and Ease of Use Questionnaire (USE) developed by Lund (2001). To measure anxiety, anxiety metrics were taken from the Unified Theory of Acceptance and Use of Technology (UTAUT) questionnaire (Venkatesh et al., 2003). In addition, six specific questions were asked about different aspects of the DSM system, to further explore the features participants did and did not like. The scoring of the USE and UTAUT were conducted by averaging the item scores in the dimension (e.g. Gao et al., 2018). All questions were administered as a Likert scale with five levels consisting of: completely disagree, disagree, neither agree nor disagree, disagree, completely disagree. See Appendix 6.1 for an overview of the questions. Finally, there was an open-ended question: do you have any other comments on the biological interactive feedback system?

### 6.3.6 Statistical analyses

All statistical analyses were performed using IBM SPSS Statistics for Windows, version 26 (IBM Corp., Armonk, N.Y., USA). Descriptive statistics were performed, and normality was verified using the Shapiro-Wilk test and visualisation of QQ plots of the unstandardized residuals. Similar to Chapter 5.3.7, $z$-scores were calculated for heart rate and skin conductance data for standardisation as no data were collected during true baseline or recovery periods (Braithwaite et al., 2012). All $t$-tests undertaken were two-tailed. Assumptions of sphericity were tested using Mauchly's test and, if significant, Greenhouse-Geisser estimates were used in the repeated measures calculations. The statistical threshold for significance was set to $p < 0.05$. Effect size was reported as partial eta squared ($\eta_p^2$) for significant results, or eta squared ($\eta^2$) for one-way ANOVAs. *Post hoc* analyses were run with Bonferroni correction.

*Self-report ratings:* For workload (NASA-TLX), a 6 (Factor: mental, physical, temporal, effort, performance, frustration) x 2 (Time: pre-, post-) repeated measures ANOVA model was run. For situation awareness, trust ratings, and reliability ratings, separate one-way repeated measures ANOVA (Journey: pre-, DSM, speed limit, post-) were undertaken.

*Physiological arousal:* A 2 (Notification: DSM, speed limit) x 2 (Time: 30 s before, 30 s after) repeated measures ANOVA was performed to understand the impact of a DSM notification on physiological arousal. The model was run for both heart rate and skin conductance level $z$-scores.

*Eye tracking:* Shapiro-Wilk test of normality and visualisation of QQ plots of the unstandardized residuals indicated that time to first fixation and duration of first fixation were not normally distributed. Data were normalised via the natural logarithm and as the response could involve a zero response, a constant (one) was added. First fixation metrics (time to first fixation and duration of first fixation) were calculated at the onset of the notification. A 2 (Notification: DSM, speed limit) x 2 (AOI: driving environment, HMI) repeated measures ANOVA was undertaken to understand whether type of notification impacted transient visual attention. For both fixation count and fixation duration, a 2 (Notification: DSM, speed limit) x 3 (Time: before, during, after) repeated measures ANOVA was calculated on the outside driving environment AOI to understand the impact of the notification on engaging visual attention.

## 6.4 Results

### 6.4.1 Self-report ratings

*Workload (NASA-TLX):* The repeated measures ANOVA model yielded a significant main effect of Time, $F_{(1, 30)} = 5.13$, $p = .031$, $\eta_p^2 = .146$; a significant main effect of Factor, $F_{(5, 150)} = 34.83$, $p < 0.001$, $\eta_p^2 = .537$; and a significant interaction of Time by Factor, $F_{(5, 150)} = 10.83$, $p < .001$, $\eta_p^2 = .265$. *Post hoc* comparisons revealed that Mental demand ($p = .003$), Physical demand ($p = .01$), and Effort ($p < .001$) decreased following simulated journeys. See Table 6.1 for descriptive statistics.

*Table 6.1.* Mean (SD) of subjective ratings for workload before and after simulated journeys.

| Subjective rating | Journey | |
|---|---|---|
| | Pre- | Post- |
| **NASA TLX [0-10]** | | |
| Mental* | 22.92 (10.64) | 16.35 (10.34) |
| Physical* | 16.47 (12.24) | 9.74 (9.86) |
| Temporal | 15.89 (13.00) | 17.73 (12.26) |
| Effort** | 24.27 (11.92) | 12.08 (10.27) |
| Performance | 23.32 (13.10) | 19.08 (12.55) |
| Frustration | 2.58 (3.91) | 5.76 (10.71) |

*Key:* * $p < .05$; ** $p < .001$

*Situation awareness (SART):* A one-way repeated measures ANOVA model showed that the main effect of Journey was not significant, $F_{(2.53, 75.86)} = 1.82$, $p = .15$. Perceived situation awareness did not differ between journeys (see Table 6.2).

*Trust and reliability ratings:* A one-way repeated measures ANOVA model showed that trust ratings significantly differed between type of Journey, $F_{(2.32, 69.61)} = 25.68$, $p < .001$, $\eta^2 = .461$. Pairwise comparisons revealed that trust ratings significantly improved from before all journeys to after all journeys ($p < .001$), before all journeys to the DSM journey ($p < .001$), and from before all journeys to the speed limit change journey ($p < .001$). There was no significant difference in trust ratings between the DSM and speed limit change journey ($p$ = .98). See Figure 6.5 for visualisation of results and Table 6.2 for descriptive statistics.

A one-way repeated-measures ANOVA model showed that reliability ratings significantly differed between type of Journey, $F_{(2,\ 60.11)} = 15.82$, $p < .001$, $\eta^2 = .345$. Pairwise comparisons revealed that reliability ratings significantly were greater following the DSM journey compared to pre-journeys ($p$ < .001), the speed limit change journey ($p$ = .03), and post-journeys ($p$ = .02). In addition, reliability ratings improved between pre- and post-journeys ($p$ < .001); and pre-journeys to the DSM journey ($p$ = .01). Overall, reliability ratings were highest following the DSM journey. See Figure 6.5 for visualisation of results and Table 6.2 for descriptive statistics.

*Table 6.2*. Mean (SD) of subjective ratings for situation awareness, trust, and reliability. Ratings provided before all journeys, after the driver state monitoring (DSM) journey, after the speed limit change journey, and following all journeys.

| Subjective rating | Journey | | | |
|---|---|---|---|---|
| | Pre- | Post- | DSM | Speed limit change |
| Situation awareness [SART] | 19.55 (5.58) | 22.35 (6.76) | 21.81 (8.28) | 20.94 (6.55) |
| Trust [0-10] | 5.06 (1.71) | 7.77 (1.71) | 8.06 (1.86) | 7.55 (2.63) |
| Reliability [0-10] | 5.45 (2.05) | 7.23 (1.69) | 8.16 (1.55) | 7.26 (2.60) |

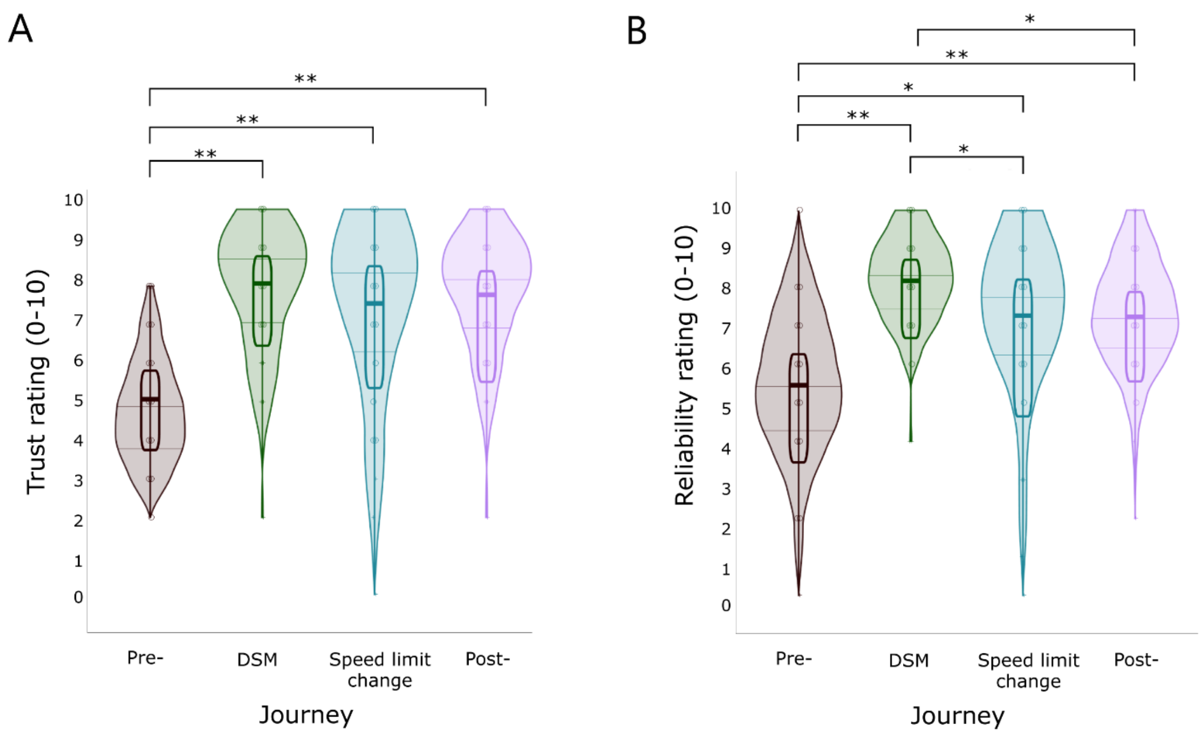



*Figure 6.5.* Violin plots representing trust and reliability ratings. **(A)** Trust ratings before all journeys, following the driver state monitoring (DSM) journey, following the speed limit change journey, post-all journeys. **(B)** Reliability ratings before all journeys, following the DSM journey, following the speed limit change journey, post-all journeys. Key: * $p < .05$; ** $p < .001$.

### 6.4.2 Physiological arousal

*Heart rate:* A two-way repeated measures ANOVA was performed to understand the impact of a DSM notification on heart rate ($z$-transformed) activity during simulated autonomous driving. The ANOVA revealed no significant main effect of Notification, $F_{(1, 27)} = 1.34$, $p = .26$; or Time, $F_{(1, 27)} = 1.99$, $p = .17$. The interaction effect was not significant, $F_{(1, 27)} = 0.20$, $p = .66$.

*Skin conductance level:* A two-way repeated measures ANOVA was performed to understand the impact of a DSM notification on skin conductance level ($z$-transformed) activity during simulated autonomous driving. The ANOVA model revealed no significant main effect of Notification, $F_{(1, 27)} = 2.27$, $p = .14$; or Time, $F_{(1, 27)} = 0.79$, $p = .38$. The interaction effect was also not significant, $F_{(1, 27)} = 0.73$, $p = .40$.

Overall, the results reveal that physiological arousal did not differ over a DSM notification or a speed limit change notification. See Table 6.3 for descriptive statistics and Figure 6.6 for a visualisation of results.

*Table 6.3.* Mean (SD) of heart rate (*z*-transformed) and skin conductance level (*z*-transformed) over notifications. Results provided for before and after a driver state monitoring (DSM) and speed limit change notification during different simulated autonomous journeys.

| Physiological indices (*z*-transformed) | Notification | | | |
|---|---|---|---|---|
| | Pre- | | Post- | |
| | DSM | Speed limit change | DSM | Speed limit change |
| Heart rate | -0.06 (0.98) | -0.72 (2.12) | -0.22 (0.80) | -0.81 (2.49) |
| Skin conductance level | -0.81 (3.07) | 1.04 (4.27) | -0.50 (2.14) | 0.95 (3.70) |

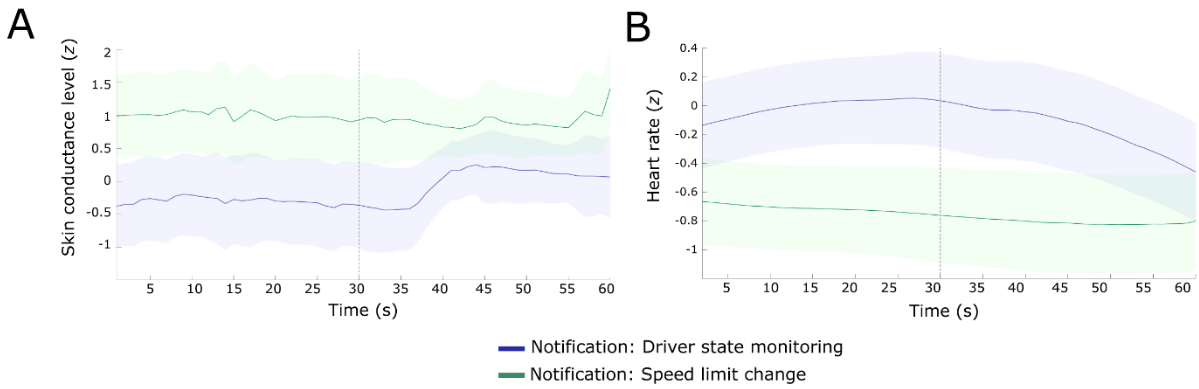



*Figure 6.6.* Heart rate and skin conductance level during a driver state monitoring (blue) and speed limit change (green) notification. Shaded areas represent the ± standard error of the mean difference. Grey dashed line represents the time point the notification appeared. **(A)** Skin conductance level (*z*-transformed). **(B)** Heart rate (*z*- transformed).

### 6.4.3 Eye tracking

*Time to first fixation:* The two-way repeated measures ANOVA revealed no significant main effects of Notification, $F_{(1, 27)} = 0.12$, $p = .74$, and AOI, $F_{(1, 27)} = 3.11$, $p = .09$, nor a significant interaction effect, $F_{(1, 27)} = 0.31$, $p = .58$. See Table 6.4 for descriptive statistics.

*Duration of first fixation:* Similarly, the two-way repeated measures ANOVA on duration of first fixation was not significant for Notification, $F_{(1, 27)} = 0.44$, $p = .51$, AOI, $F_{(1, 27)} = 0.26$, $p = .61$, nor the interaction, $F_{(1, 27)} = 1.87$, $p = .18$. See Table 6.4 for descriptive statistics representing first fixation metrics.

*Fixation count:* The two-way repeated measures ANOVA model yielded a significant main effect of Time, $F_{(2, 54)} = 37.62$, $p < .001$, $\eta_p^2 = 0.58$, Notification, $F_{(1, 27)} = 23.36$, $p < .001$, $\eta_p^2 = .51$, and an interaction effect, $F_{(2, 54)} = 5.22$, $p = .008$, $\eta_p^2 = 0.16$. Pairwise comparisons revealed that fixation count was always greater during the DSM message when compared to the speed limit change message, before ($p = .016$), during ($p < .001$) and after ($p = .03$) the notification. During the DSM message, fixation count was greater before compared to during ($p < .001$), and after compared to during ($p < .001$). Similarly, for the speed limit change message, fixation count was greater before compared to during ($p <. 001$) and after compared to during ($p < .001$). See Table 6.5 for descriptive statistics.

*Fixation duration:* The two-way repeated measures ANOVA model revealed a significant main effect of Time, $F_{(2, 54)} = 12.14$, $p < .001$, $\eta_p^2 = .31$, and Notification, $F_{(1, 27)} = 7.028$, $p < .013$, $\eta_p^2 = .21$. The interaction was not significant, $F_{(2, 54)} = .36$, $p = .70$. Pairwise comparisons revealed that fixation duration was greater during the DSM message ($M = 57.16$, $SD = 15.56$) when compared to the speed limit change message ($M = 48.61$, $SD = 11.79$), regardless of time ($p = .013$). In addition, fixation duration was lower during the message ($M = 43.51$, $SD = 12.76$) when compared to before the message ($M = 56.95$, $SD = 16.65$; $p = .002$), and after the message ($M = 58.19$, $SD = 14.78$, $p < .001$). See Table 6.5 for descriptive statistics.

Overall, the eye tracking results suggest that neither message impacted first fixation more so than the other, and fixation count and duration were greater during the DSM notification, regardless of time. During the notification period, fixation duration and count were lower. See Figure 6.7 for a visualisation of results.

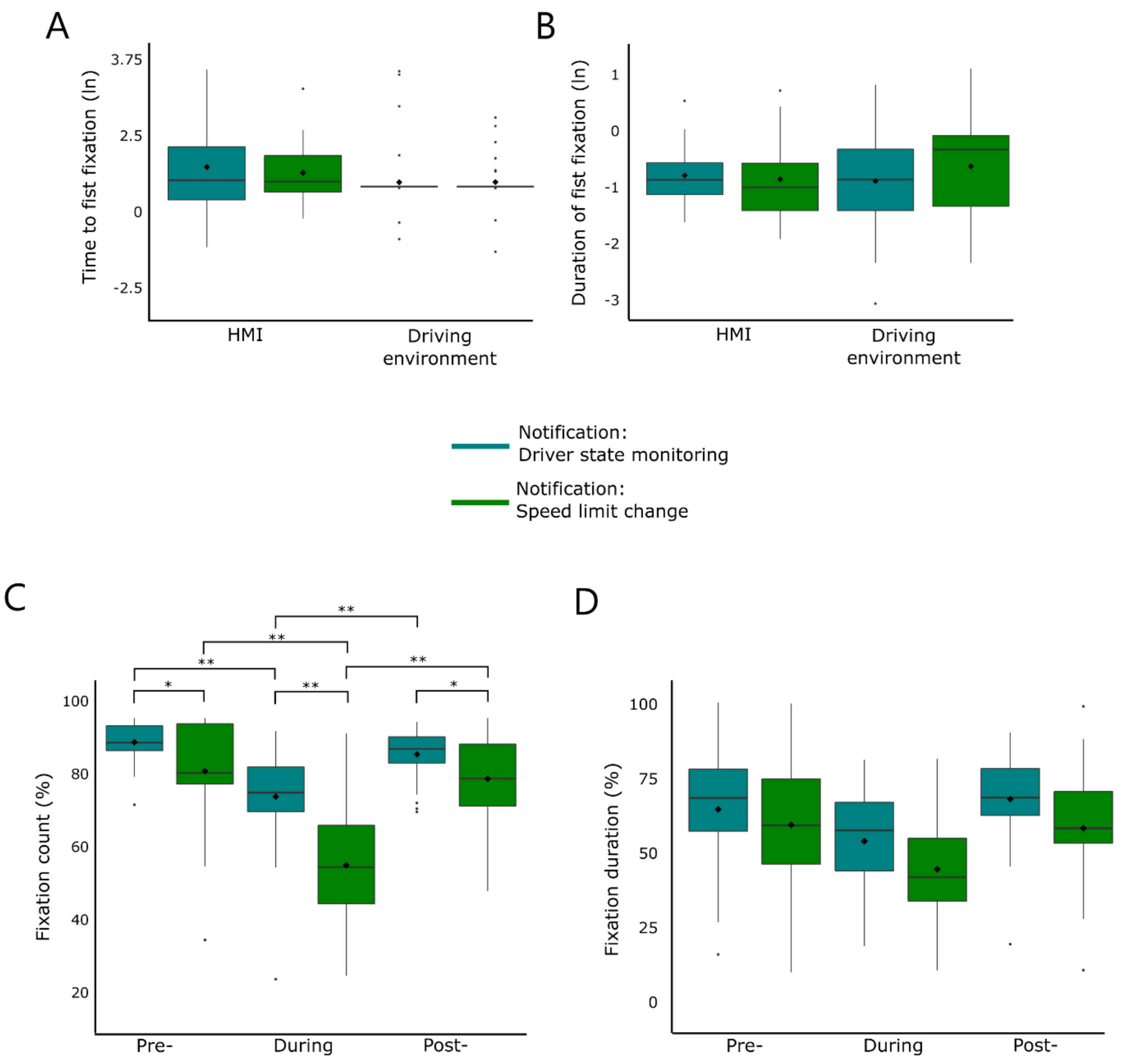



*Figure 6.7.* Fixation metrics during a driver state monitoring notification journey and a speed limit change notification journey. **(A)** Time to first fixation (ln) during the notification on different areas of interest (AOI). **(B)** Duration of first fixation (ln) during the notification on different AOIs. **(C)** Fixation count (%) on the driving environment before, during, and after the notification. **(D)** Fixation duration (%) on the driving environment before, during, and after the notification. Key: ** $p < .001$; * $p < .05$; ⬩= mean value.

*Table 6.4.* Mean (SD) of natural logarithm transformed fixation metrics during a driver state monitoring (DSM) and speed limit change notification in Areas of Interest (AOI): HMI and driving environment.

| Fixation metric (ln+1) | AOI x MESSAGE | | | |
|---|---|---|---|---|
| | HMI | | Driving environment | |
| | DSM | Speed limit change | DSM | Speed limit change |
| Time to first fixation | 1.58 (1.40) | 1.41 (0.84) | 1.13 (1.26) | 1.14 (0.71) |
| First fixation duration | -0.73 (0.47) | -0.79 (0.68) | -0.82 (0.86) | -0.57 (0.90) |

*Table 6.5.* Mean (SD) of fixation metrics (%) pre-, during, and post- a driver state monitoring (DSM) and speed limit change notification.

| Fixation metric (%) | TIME x MESSAGE | | | | | |
|---|---|---|---|---|---|---|
| | Pre- | | During | | Post- | |
| | DSM | Speed limit change | DSM | Speed limit change | DSM | Speed limit change |
| Fixation duration | 59.70 (19.83) | 54.21 (23.29) | 48.46 (18.24) | 38.56 (17.23) | 63.33 (16.88) | 53.06 (19.26) |
| Fixation count | 91.68 (7.03) | 81.65 (18.21) | 72.84 (16.41) | 49.06 (20.85) | 87.46 (8.95) | 79.02 (15.31) |

### 6.4.4 Subjective experience of the driver state monitoring system

The DSM system received an average usefulness rating of 3.78 (*SD* = 1.04). Overall, the system was rated positively, with 65% of responses being either agree or strongly agree. Only 3% of responses were disagree, and no responses were rated as strongly disagree. In regard to anxiety, the system received an average rating of 2.60 (*SD* = 1.22) suggesting that the system was neither anxiety-provoking nor not anxiety-provoking. 20% responses were rated as strongly agree or agree, suggesting some anxiety. 39% responses were strongly disagree or disagree. See Figure 6.8 for visualisation of results.

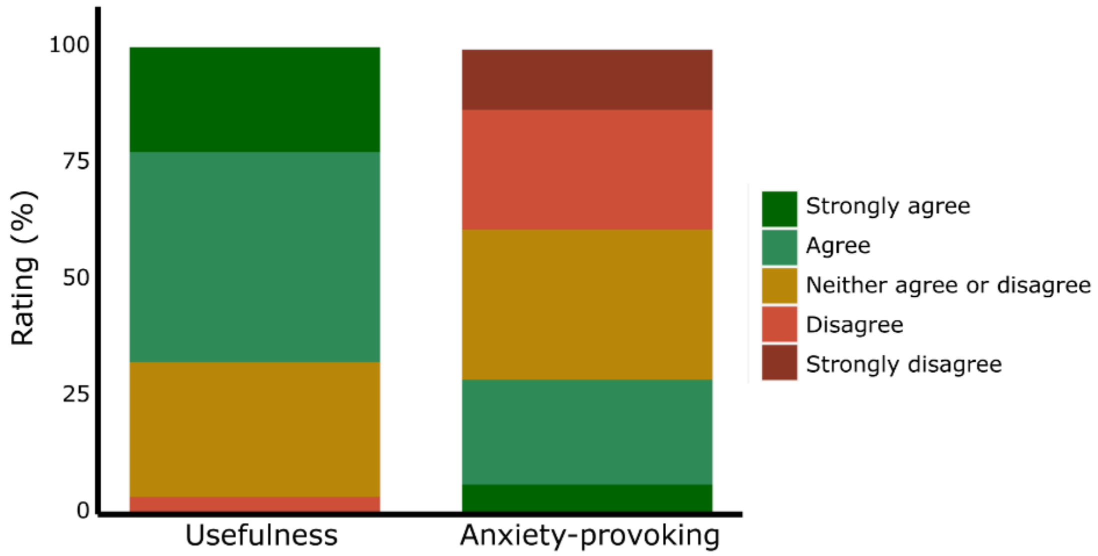



*Figure 6.8.* Likert-style response to usefulness and anxious dimensions regarding the driver state monitoring system.

Next, the six Likert-style questions regarding specific aspects of the driver state monitoring system were individually analysed. See Figure 6.9 for a visualisation of the results. Participants expressed that they would want to be able to turn the system on and off, with an average rating of 3.94 (*SD* = 1.01). 71% of responses were either strongly agree or agree, with only 13% of responses being disagree and no responses rated as strongly disagree. Participants averaged rating was 2.48 (*SD* = 1.04) to the question "I liked that the system took control and changed the speed of the vehicle". Most participants disagreed or strongly disagreed with this comment (61%) whereas only 25.81% of participants agreed, and no participants strongly agreed. On the other hand, 55% of participants either strongly agreed or agreed to the question "I would like to control and change the speed of the vehicle myself". 23% disagreed or strongly disagreed, and the average rating was 3.45 (*SD* = 1.07). Participants wanted to have options if the driving style was going to change. 93% of participants either agreed or strongly agreed to this question, while no one disagreed or strongly disagreed. The average rating was 4.53 (*SD* = 0.62). Participants also expressed that they wanted to know how the system worked. This question received an average rating of 4.06 (*SD* = 1.11), with 81% of participants agreeing or strongly agreeing, compared to 10% disagreeing or strongly disagreeing. Finally, participants liked that the system alerted them via the HMI dashboard. The average rating was 4.58 (*SD* = 0.55), suggesting explainability and HMI interaction is paramount. 97% of participants strongly agreed or agreed with this question, with no participants disagreeing or strongly disagreeing.

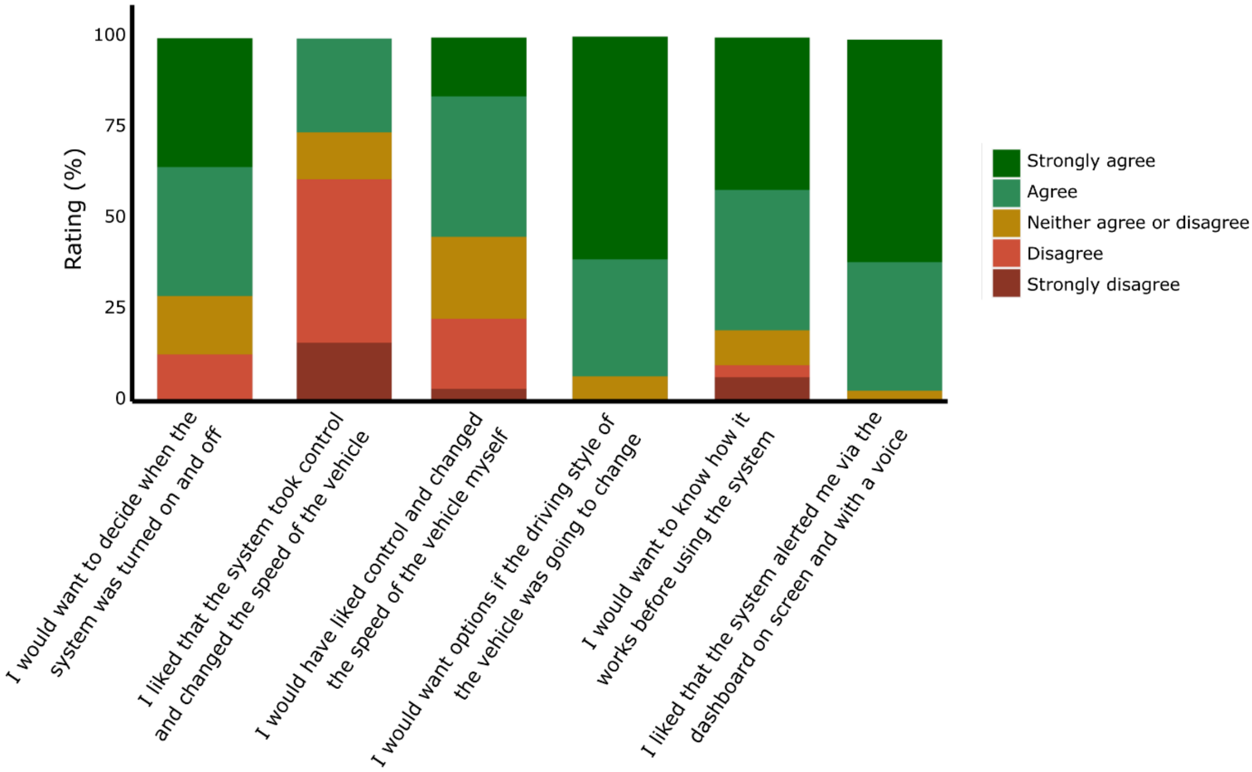



*Figure 6.9.* Likert-style responses regarding the driver state monitoring system.

15 participants (48%) responded to the open-ended question asking if they had any further views on the driver state monitoring system. Overall, the majority of responses fell into two categories: the system improved safety and driver attention (five responses), and the system needed more user control (five responses). The other five responses consisted of different aspects of the experience that could be improved. Two responses suggested that the vehicle changed speed for too long which could be dangerous. Another response indicated that if the vehicle changed behaviour automatically it could cause frustration. One response was concerned with the potential anxious feelings it could provoke (e.g. why has my heart rate increased?). The final response suggested the system was limited as it only considered physiological response.

Overall, the Likert-style question and the open-ended question demonstrate that an element of user control and clear notifications from the driver state monitoring system is paramount for user acceptance.

### 6.5 Discussion

This study aimed to understand the impact of a driver state monitoring (DSM) notification on ocular attention, arousal, and user acceptability. Eye gaze and physiological data between different simulated autonomous journeys consisting of a DSM notification and a speed limit change notification were compared. The results revealed several findings. Perceived situation awareness did not differ, but mental demand, physical demand, and effort factors decreased following the simulated autonomous journeys compared to pre-values. Trust ratings improved following the journeys but did not differ following the DSM or speed limit change journey. Yet, reliability ratings did significantly improve following the DSM journey, when compared to the speed limit change journey. Physiological arousal was not impacted by either notification. Eye gaze metrics revealed that fixation duration and count was significantly greater before, during, and after a DSM notification, when compared to a speed limit change notification. Moreover, participants found the DSM system useful, although expressed a need for more input to improve the safety and usability of the system.

#### 6.5.1 Following all autonomous journeys, physical and mental demand and effort decreased, yet situation awareness did not change

Comparable to Chapter 5, physical demand and effort workload factors decreased after all simulated journeys. This provides support that these workload factors are not recruited during automated driving regardless of the environment (simulated versus real world). In addition, mental demand decreased following autonomous journeys. This is distinct from Chapter 5, as mental demand did not decrease during real-world autonomous journeys, suggesting a simulated environment requires less demand on cognitive processes, which may reflect the higher risk in real-world driving, or the more complex visual environment. This contradicts previous research that has found that the mental demand factor is greater during simulated driving when compared to real-world driving (Galante et al., 2018). However, Galante et al.'s (2018) findings may be due to a lack of familiarity with a driving simulator. Steering and acceleration manoeuvres can at times feel unrealistic (Kharrazi et al., 2020) and often extended periods of time (e.g. 20 minutes) are required to adapt. During autonomous driving, perceptuo-motor calibration is less important, which might be why in the present study mental demand ratings decreased.

Simulated autonomous journeys and type of notification had no impact on perceived situation awareness. De Winter et al. (2014) meta-analysis and review concluded that situation awareness increased if the participant is notified or instructed through a human-machine interface (HMI). During the present study, the notifications were presented on an HMI, yet participants did not report greater situation awareness in either journey. In addition, the eye tracking results revealed that participants were more likely to gaze at the roadway during the DSM journey, which is often coupled with an increase in subjective situation awareness (De Winter et al., 2014). However, there are clear discrepancies between this study and the ones reported by De Winter et al. (2014). Firstly, the studies reviewed compared either automated driving with manual driving, or automated driving with adaptive cruise control. Therefore, differences between the notifications may not have been detected due to the similar operational, tactile, and strategic tasks employed during both autonomous journeys. Secondly, the authors did not include studies that utilised self-report measures of perceived situation awareness, including the SART. As such, it is not accurate to compare the present study results with De Winter et al.'s (2014) findings. Previous research has also criticised the face and construct validity of the SART. Salmon et al. (2009) found that the scale did not correlate with performance or other measures of situation awareness. They also highlighted that the scale focuses on attention, complexity, and variability of the situation, which are related to workload rather than situation awareness. As automated driving arguably decreases workload factors (as evidenced by the NASA-TLX results), this may have inhibited any variations in perceived situation awareness.

### 6.5.2 Trust increased following all journeys, but reliability was highest following the driver state monitoring journey

Self-report trust ratings increased following all autonomous journeys, which likely reflects participants' positive evaluation of the autonomous vehicle over journey experiences (i.e. learned trust, Hoff & Bashir, 2015). Interestingly, reliability ratings were greater following the DSM journey, when compared to the speed limit change journey. Although trust and reliability are arguably intermixed (e.g. Akash et al., 2018), these results suggest that the vehicle was most dependable during DSM. Both notifications informed participants the reason for its behaviour (i.e. vehicle slowing down due to driver state/speed limit change), yet the DSM notification offered another level of technical competence by monitoring driver state, a likely novel concept for the participants. Thus, the system transparency and/or the uniqueness of the DSM message may have encouraged higher reliability ratings.

These results have important implications for DSM. Reliability, like trust, endures a complex relationship with automation misuse. Too much dependability may lead to accepting automation errors, whereas too little dependability may lead to disregarding the system when it might provide a benefit (McBride et al., 2011). The results therefore suggest that reliable DSM systems may encourage sub-optimal monitoring of the automated vehicle or reluctance to counteract the system, due to over-complacency. Varying reliability during automation monitoring has been found to increase detection of automation errors (Prinzel et al., 2005) and may therefore be able to disrupt over-complacency.

### 6.5.3 Physiological arousal was not modulated by notifications

Heart rate and skin conductance level were similar before and after both types of notifications, indicating that a DSM notification had no impact on physiological arousal indices. The results demonstrate that simply telling participants that their arousal indices indicate performance decrements does not instigate modulation of autonomic arousal.

Visualisation of the results demonstrate increased skin conductance level approximately five seconds following a DSM notification (see Figure 6.6). This response is similar to the skin conductance response in Chapter 5, indicating increased sympathetic arousal following a DSM notification. Despite this increase, skin conductance was relatively low when compared to a speed limit change notification, including before the message appeared. This pre-period acted as a condition-specific baseline period and therefore it is not clear why autonomic arousal was lower during DSM. To investigate whether wrist movements may have confounded the measurement period, including the baseline, movement index was calculated between the times of interest (see Appendix 6.2). The results revealed that wrist movement was significantly greater during the speed limit change notification, compared to the DSM notification. Therefore, it is possible that wrist movements, potentially from using the HMI, may have confounded physiological data, resulting in enlarged values during the speed limit change notification.

Despite indices such as heart rate and skin conductance to stimuli being extensively studied, the relation between HMI notifications and physiological signals remain largely unexplored. A recent study demonstrated reduced heart rate and skin conductance levels during a simulated autonomous journey containing visual and audio notifications when compared to notifications without audio (Eimontaite et al., 2020). The authors suggest that increased system transparency reduced physiological arousal. However, the authors focused on physiological signals during the whole journey, which would have consisted of different driving scenarios and other human-vehicle interactions. In this study, real-time physiological response to the notification were analysed, and so only focused on two 30 second periods throughout the journey. Research investigating other types of notifications can provide further insight. For example, Fortin et al. (2019) found that smartphone notifications increased skin conductance responses. Yet, smartphone notifications have a social element, which is largely known to modulate arousal mechanisms. For example, social evaluative threat increases sympathetic and parasympathetic activity (Dickerson & Kemeny, 2004). In the present study, notifications were related to the driving task, and did not include a social aspect. This might partly be why arousal did not increase at the onset of either notification.

Overall, the physiological results are inconclusive. The statistics indicate that a DSM message did not result in adverse effects as characterised by over-activation or inhibition of autonomic arousal indexed by heart rate and skin conductance, yet wrist movements may have confounded the comparisons.

### 6.5.4 Eye gaze was not modulated by notifications

The eye fixation results revealed that the DSM notification was no more attentionally demanding compared to a speed limit notification. This is evidenced by the similar time to first fixation and first fixation duration metrics between messages. Secondly, fixation count and duration were greater on the driving environment during a DSM journey, compared to a speed limit notification journey. However, this was apparent before, during, and after the notification.

The similar first fixation metrics between notifications and areas of interest suggest that auditory and verbal notifications do not distract the passive driver away from the driving environment. Despite the uniqueness of a DSM notification, the lack of significant effects for duration of first fixation suggests that the notifications had no distinct impact on focused attentional processing mechanisms. First fixation metrics provide a reflection of early object identification processing (Henderson et al., 1987) however others argue that they reflect perceptuo-motor factors such as foveal landing position rather than cognitive processes (e.g. Arizpe et al., 2012). Despite this, first fixation metrics can provide insight into when participants decided to direct their gaze towards the notification. Research has indicated that eye gaze is more likely to be directed towards novel (e.g. Horstmann & Herwig, 2016) and salient stimuli (e.g. Giraudet et al., 2015). Despite this, participants did not fixate for longer on the DSM notification. The present study controlled for differences between the physical properties of the notifications so that any potential difference between fixation metrics could not be associated with bottom-up processes only (e.g. same contrast, font size, colour etc). This may partly explain why there was no significant difference between notifications. Appearance properties such as location (Ozimek et al., 2019), opacity (Imbert et al., 2014), size and motion dynamics (Athènes et al., 2000) impact how distracting a notification is (Klauck et al., 2017).

However, interpreting first fixation metrics in this manner assumes that cognitive processing of the message finishes once gaze is directed away from the HMI. But information processing can still be ongoing once fixations ends (Anderson et al., 2004). Therefore, to try and understand whether visual attention was encouraged towards the roadway, gaze duration and count metrics on the driving environment was compared before, during, and after the notification. Overall, these results were inconclusive. Although fixation count and duration were greater during DSM, this effect was apparent before the message appeared, during what could be considered a baseline period. It is difficult to ascertain why these differences before the notification appeared, however it may be attributed to the journey instructions or the simulated scenario. Before the DSM journey, participants were reminded to look out for obstacles and press the 'safe stop' icon on the HMI if they wanted to stop the vehicle. Although this was also mentioned at the beginning of the testing session, participants were reminded to ensure that they were clear about why monitoring of physiological state was important: in case their performance (i.e. ability to press the safe stop button) was delayed during autonomous driving. However, during the journey, participants were not required to press the safe stop button, nor did they experience any type of obstacle. Alternatively, fixation metrics before the notification may have been impacted by the simulated environment. Although journeys were randomised, the DSM message and speed limit change message were displayed at the same location during the route. Before the DSM message, the participants experienced a 'Town centre' environment. The visual scene complexity may have engaged visual gaze (Pomplun et al., 2013) in comparison to the less built up environment which appeared before the speed limit change message. Yet this is only speculation, and so it is not clear whether visual attention was unaffected by DSM, whether it was covered by a ceiling effect, or whether it was impacted by task instructions.

#### 6.5.5 The driver state monitoring system was perceived as useful, but users wanted more input and control

At first glance, the questionnaire results indicate that DSM is a valuable contribution to an advanced driver assist system. Participants expressed alerting the driver was useful and could increase self-reported attention. In line with Melnicuk et al. (2019), the overall acceptance of a DSM system was impacted by usefulness metrics, particularly whether it was safe and easy to use. In addition, in line with the increased reliability ratings following DSM, Melnicuk et al. (2019) found that reliability was an important factor concerning the overall usability of a DSM system. However, there was an emphasis for more user input, such as the ability to turn the system off, or to ask the driver how they wanted the vehicle to respond. Participants also expressed that this could lead to frustration, and even anxiety. In some cases, the vehicle's behaviour could be dangerous - reducing vehicle speed would not be appropriate for all road conditions. Therefore, a DSM system could have a negative impact on driver wellbeing and comfort. These results echo similar studies on technology and user acceptance, particularly autonomous features of a vehicle. For example, Choi & Ji (2015) suggest system transparency and technical competence are important factors that enable trust. Yet, the results are distinct to those found by Melnicuk et al. (2019). Overall, the authors found that anxiety did not have an impact on participants' intention of using a DSM system. However, their participants did not experience a DSM system, and only answered questions relating to their willingness to use it. The present findings therefore add to their findings: acceptance behaviour may be impacted by anxious feelings. Overall, these results provide additional insight for future research studies. Participants who experienced a DSM system found it useful, however an on-going challenge concerns how information is best delivered, and what human input is appropriate for a DSM system to improve safety and wellbeing.

### 6.5.6 Limitations

The present study did not instigate an unexpected event and so did not measure the impact of a notification on deficit attentional strategies during an unexpected event. Chapter 5 demonstrated the relationship between a vehicle hazard and visual attention deficits, and so it would be logical to next assess whether a DSM system could promote suitable attentional strategies in response to a critical event. This might also explain why the present study did not uncover any differences between notifications in autonomic arousal and visual attention. However, due to the limited research in the area, this study was conducted to understand the positive or negative impact of a DSM *message* on driver state. The aim of the study was to understand how the message itself was perceived and the behaviours associated with it. This is important, as inappropriate implementation of DSM systems could have damaging consequences to the acceptability of such systems, as well as promote unsafe human-vehicle behaviours (e.g. overreliance).

To ensure participants were not confused when presented with the DSM notification, participants were told that their physiological signals were being monitored by a real-time system during the journeys. Although the Empatica E4 did collect participants' heart rate and skin conductance level, this was not being fed into a real-time system, and so all participants received the pre-programmed DSM notification at the same location during the journey, no matter what their physiological signals read. Yet, participants may have had interoceptive awareness: the ability to sense internal state, including heart rate or skin conductance. As such, participants may have felt that their heart rate or skin conductance had not changed and so dismissed the notification. Despite this, all participants provided their subjective experience of the DSM system and no one indicated any concerns with the physiological measures or mentioned that the notification was not in alignment with their own perceptions at the time.

### 6.5.7 Conclusion

Monitoring of driver state has the potential to improve safe human-vehicle interaction during semi-autonomous driving and the development of fully autonomous vehicles, as well as wellbeing of passengers during autonomous driving. Although there is a surge of engineering research examining new remote and wireless systems, there is a lack of understanding of the impact on DSM on cognitive demand and general wellbeing. Therefore, it is not clear if DSM notifications will have the reverse anticipated effect and exacerbate attentional deficits. In addition, it is not clear what the potential barriers are for user acceptance. This simulated autonomous driving study investigated ocular attention and physiological arousal during two types of notifications which preceded the vehicle slowing down: a DSM message of changes in physiological state, and a conventional message of a speed limit change. Participants rated the reliability of the vehicle the highest after a DSM journey. There were no differences in the real-time autonomic response between notifications. First fixation metrics revealed that both notifications involved similar visual processing demands. Participants made more fixations during a DSM notification when compared to a speed limit change notification, yet this difference was also apparent before the notification, indicating that driving context may have been responsible for this difference. Moreover, participants found the DSM system useful, although expressed a need for more input to improve the safety and usability of the system. Overall, the Chapter found that a DSM notification reporting on physiological indices was not visually distracting and did not modulate physiological arousal compared to other driving-related notifications. Therefore, a DSM system could be a useful adjunct to autonomous vehicles. As Chapter 5 revealed potential negative experiences during Level 5 driving, coupled with human limitations in sustained monitoring during high arousal situations, further research is needed to determine whether a DSM notification successfully engages attention mechanisms during periods of low or high arousal, such as unexpected situations.

# 7.0 General Discussion

## 7.1 Summary of thesis

Sustained attention is a multifaceted concept that operates on a series of timescales, impacted by global mechanisms of both vigilance and arousal, as well as being a key element to many higher order cognitive functions (Langner & Eickhoff, 2013). Decrements in attention can be disastrous in occupational and everyday settings that require sustained monitoring, which is becoming more and more prevalent with the rapid evolution of robotic and autonomous systems. Although these systems have the potential to enrich environments, safe and efficient human-machine interaction is challenged by the negative effects on cognition when monitoring automation and intervening when automation fails (Parasuraman & Riley, 1997). Previous research has indicated that autonomous systems encourage cognitive underload (e.g. Carsten et al., 2012; Louw & Merat, 2017), reduce spare attentional capacity (e.g. Stapel et al., 2019), as well as modulate social cognitive factors such as trust (e.g. Körber et al., 2018; Walker et al., 2019). However, with new autonomous systems developing, what still eludes researchers is a complete understanding of attentional lapses during continuous tasks, and how the engagement of a previous unrelated task can impact current task processes and performance. In addition, it is not clear whether attention fluctuations can be captured reliably during naturalistic settings including autonomous driving. Considering that interactions within the real world are endlessly evolving, the temporal dynamics of brain oscillations, as well as more slower moving autonomic processes, may provide complementary evidence into how cognitive processes such as attention are affected by these continuous interactions.

This thesis aimed to investigate variations in arousal and attention, as well as performance fluctuations, during continuous tasks. Specifically, this thesis focused on an autonomous vehicle environment. Across the studies, physiological measures of skin conductance, heart rate, heart rate variability, muscle activity, fronto-parietal alpha and theta neural oscillations were explored. Behavioural measures of task performance and eye tracking, as well as subjective ratings, were collected when appropriate. In five empirical studies I aimed to understand the lasting effects of previous tasks and how attentional demand impacted autonomic arousal and neural mechanisms related to sustained attention (Chapters 2, 3, and 4). I then investigated how attention fluctuations manifested, and the feasibility of capturing these fluctuations during real-world monitoring of automation during an unexpected (i.e. high arousal) event (Chapter 5). Finally, I attempted to understand whether continual psychophysiological monitoring is a feasible and acceptable approach to modulate behaviour and wellbeing (Chapter 6). I used both controlled laboratory-based and applied experiments to unpick the research questions in an attempt to improve the translation of research findings to human state monitoring in naturalistic settings.

This final chapter is arranged in the following manner. The first section will focus on the findings from the laboratory-based experiments on autonomic and neurobehavioural evidence for attentional deficits during continuous multitasking (Chapter 2), and the impact of a previous task differing in attentional demand on visuomotor performance (Chapter 3) and simulated driving (Chapter 4). The second section focuses on the findings from the applied autonomous vehicle 'Pod' study (Chapter 5) which was undertaken to reveal how ocular behaviours and autonomic arousal was impacted by an unexpected situation during Level 5 autonomous driving; and the virtual driving simulator study (Chapter 6) to explore the attentional demand and user acceptability of driver state monitoring notifications, during Level 5 autonomous driving. Given the applied aspect of this research, and the argued potential benefits for psychophysiological human monitoring, the present Chapter will conclude by discussing challenges associated with the findings of the thesis, before discussing future directions.

## 7.2 Summary of findings

### 7.2.1 Neurobehavioural evidence for performance deficits during multitasking: the impact of previous task demands (Chapter 2, 3 and 4)

#### *7.2.1.1 Overview*

The first experiment described in Chapter 2 aimed to investigate whether various subjective, physiological, and neurophysiological measures were sensitive to the influence of attentional load on multitasking. Attentional load was indexed by workload intensity (low versus high) and induced via the Multitasking Framework, a demanding computer task that required the participant to attend and respond to four tasks simultaneously. Although physiological and neurophysiological measures revealed some variation in response to workload conditions, only subjective ratings were able to successfully disentangle workload, and therefore indicated a high sensitivity of subjective ratings to workload changes during multitasking. Neural oscillatory activity revealed increased fronto-parietal theta power during multitasking, reflecting task engagement, and this persisted following task cessation. Alpha power increased only following task cessation, potentially indicating a decrease in general arousal processes. The findings therefore revealed a lasting neural effect, specifically fronto-parietal theta power, on multitasking. As the results were limited, Chapter 3 and 4 attempted to unravel the negative lasting impacts of multitasking by employing an event-related design to increase the signal-to-noise ratio and remove within-subject variation, to separate workload conditions.

In Chapter 3, an event-related design was employed to explore whether the attentional load of a previous task impacted the general mechanisms of visuomotor performance. To this end, prior attentional load (low versus high visual search) was manipulated, while keeping the physical properties of the secondary task - visuomotor pursuit tracking - the same. Although reaction time measures were similar following both tasks of varying attentional load, participants spent longer outside of the target (total duration of deviations) following high load, reflecting impaired performance. Short-lasting bursts of theta synchronisation (ERS) and continuous alpha desynchronisation (ERD) were most pronounced at frontal and parietal scalp locations, respectively. Following a visual search task of high load, frontal theta power was delayed approximately 100ms, possibly reflecting delayed target encoding mechanisms. In addition, contralateral alpha desynchronisation was greater (i.e. lower alpha power), during the first and last second (approximately) of the continuous tracking task. These results suggest that demands on attention were greater following engagement in a high load visual search task.

These initial results suggest a lasting effect of attentional demand when switching to a new task. Therefore, a similar method was employed in Chapter 4 to explore whether engagement with a previous task (passive viewing, low load visual search, high load visual search) had an impact on takeover during simulated semi-autonomous driving. The passive viewing condition was considered a control condition, where participants were required to watch the roadway instead of engaging in an unrelated task. Both takeover time and takeover quality measures were negatively impacted by engagement of a visual search task, regardless of attentional load. Firstly, time to accelerator pedal was greater following a visual search task of low load, and a task of high load, when compared to the passive viewing task, indicating an initial delayed ability to regain awareness of the vehicle controls. Standard deviation of lateral position was greater following a visual search task of low load, and a task of high load, when compared to the passive viewing task. This suggests that participants demonstrated less control of the vehicle, ‘weaving’ in and out of the lane. Similarly, mean vehicle speed was lower (deviated further from the 60 mph speed limit) following a task of low load and following a task of high load, when compared to passive viewing. This could be due to allocation of cognitive attentional resources to taking over, and therefore the delayed ability to regain general situation awareness including monitoring speed. Neural evidence supports this hypothesis, as frontal theta synchronisation was seemingly absent following the passive viewing task, representing the effortlessness of re-allocating external attention processes. Attention-related parieto-occipital alpha power was less pronounced (lower desynchronisation) following the task of high load when compared to the passive viewing task, suggesting deficits in attention allocation. Furthermore, alpha power negatively correlated with vehicle speed, demonstrating that higher alpha power (lower desynchronisation) was associated with better performance. Altogether, the results suggest that engaging in unrelated tasks of high attentional load before a takeover request may result in deficits in attentional processes, as indexed by parieto-occipital alpha, while negatively impacting driving performance. See Table 7.1 for an overview of behavioural and EEG results of interest.

*Table 7.1.* Overview of behavioural and EEG results of interest in Chapter 3 and 4.

| Measure | Tasks employed | |
|---|---|---|
| | Chapter 3<br>Visual search low load, visual search high load | Chapter 4<br>Passive viewing, visual search low load, visual search high load |
| **Behavioural** | | |
| Instant measures of performance (i.e. reaction times) | No differences following a task of low or high load | Time to accelerator pedal was greater following low and high load task compared to passive viewing |
| Continuous measures of performance (i.e. accuracy) | Duration of deviations (time outside the target) was greater following high load compared to low load | Speed deviation and standard deviation of lateral position was greater following low and high load task compared to passive viewing |
| **EEG** | | |
| Frontal theta ERS (dB normalised) | Delayed following a task of high load when compared to low load | Greater following a task of low and high load when compared to passive viewing |
| Parieto-occipital alpha ERD (dB normalised) | - | Greater following passive viewing when compared to high load |
| Contralateral parietal alpha ERD (dB normalised) | Greater following high load when compared to low load | - |

*Key:* ERD = event related desynchronisation; ERS = event related synchronisation; - = not measured

#### *7.2.1.2 Discussion*

Chapter 2 revealed a lasting effect of increased frontal theta following multitasking (regardless of load), while Chapters 3 and 4 revealed somewhat distinct results. In Chapter 3, a transient burst of frontal theta activity was delayed approximately 100 ms when following a task of high attentional load, while Chapter 4 revealed no significant differences when following a task of low or high load, however, theta was absent following the control passive viewing task. Overall, these results are in accordance with the literature. The selection of new information reliably elicits a strong event-related increase in fronto-midline theta (Klimesch, 2012). For example, a stronger increase in theta is seen during the retention of six items, when compared to two items (Jensen et al., 2002). As such, theta was absent when no new information was provided (e.g. passive viewing condition in Chapter 4). However, the onset of theta differed between Chapters 3 and 4. Theta commenced at approximately a similar time (200 ms post-cue) following a task of high load in Chapter 3 and following a task of low and high load in Chapter 4. However, following a task of low load in Chapter 3, theta commenced at 100 ms, suggesting rapid processing of new information. As the visual search task was the same between Chapter 3 and 4, the results suggest that the difficulty of the continuous task (i.e. tracking versus simulated driving), as well as the attentional load of the previous task, impacted the time course of frontal midline theta. Theta appeared earlier following visual search low load during the tracking task and appeared later following visual search low load during simulated driving.

An important distinction between Chapter 3 and 4 is the spatial distribution of attention-related alpha during both visuomotor tasks. Although alpha desynchronisation was bilateral in both Chapters, variations in alpha desynchronisation dominated the left hemisphere during the tracking task, whereas alpha changes were bilateral during simulated driving. The varying demands between the two tasks may explain these differences. During the simplistic tracking task, the increase in left parietal alpha desynchronisation during high load may represent a stronger recruitment of motor attention processes (Rushworth et al., 2001). Mechanisms were recruited to redirect the intended (right hand and arm) movements, as potentially, there were not enough resources available for accurate movement selection. By contrast, driving is a complex visuomotor task and changes in alpha are reliably elicited at bilateral parieto-occipital electrodes during simulated driving (e.g. Chuang et al., 2018; Ma et al., 2018; Wang et al., 2018). Brooks et al. (2016) found that alpha in the right parietal area was related to driving metrics associated with orienting the vehicle on the road, measured via lane deviation; whereas left parietal alpha activity was related to corrective steering responses and heading error. The authors argue that alpha in the right parietal cortex was related to processes which monitor vehicle kinematics, whereas alpha in the left parietal cortex reflected feedback control processes such as corrective steering responses. Although alpha was only examined during a pre-perturbation period (when the vehicle would veer left or right somewhat mimicking a gust of wind), these results provide insight into the fluctuations in alpha likely reflecting dynamic attention allocation during driving, which could have been impacted by a previous task of varying attentional load. In sum, varying attentional demands of the previous task potentially only interfered with motor attention during the simplistic tracking task, whereas disruptions to further processes such as orienting may have arisen during simulated driving.

A further distinction in attention-related alpha can be found between Chapters 3 and 4. At first glance, the alpha results seems contradictory. Chapter 3 revealed greater alpha desynchronisation following high load reflecting increased demands on attention processes. This finding was expected, as accentuated alpha desynchronisation during a demanding task driven by sensory inputs is most commonly described in the literature (e.g. Magosso et al., 2019; Proskovec et al., 2019; Sauseng et al., 2006; see Borghini et al., 2014 for a review). However, Chapter 4 revealed *reduced* alpha desynchronisation following high load. If attentional demand was increased following a task of high load in both Chapters 3 and 4, why did alpha power changes manifest in opposite directions during the tracking task and the simulated driving task? Chapter 4 can provide some insight as the passive viewing task acted as a control condition, as participants did not take part in a visual search task but monitored the roadway ahead. Following passive viewing, alpha desynchronisation was greater, when compared to following the visual search task of high load. In addition, a negative correlation between vehicle speed and parieto-occipital alpha revealed that increased alpha desynchronisation (i.e. lower alpha power) was associated with better performance. In combination, the results from Chapter 3 and 4 suggest that reduced alpha desynchronisation (i.e. higher alpha power) was associated with attentional deficits. Therefore, it is likely that the prior task of high load increased demands on resources for the following continuous task. However, as the driving task was more demanding than the tracking task, not enough resources were available for successful task performance. As such, alpha desynchronisation was lower during high load driving. Although somewhat speculative, this interpretation is also supported by the literature, which has demonstrated that attenuated alpha desynchronisation is related to poorer performance (e.g. Dimitrijevic et al., 2017; Proskovec et al., 2019; Puma et al., 2018). However, the results are somewhat inconclusive, as although Chapter 4 revealed a significant difference between low and high load, Chapter 4 did not. This may have been due to the differences in the 'task switch' between the visual search task and the continuous performance task (i.e. tracking/simulated driving). For Chapter 3, participants engaged with one computer screen. In Chapter 4, participants were required to shift their visual gaze and body position from a computer screen displaying the visual search, towards the simulated driving environment. This move to a new location would have required additional perceptual and motor processes, resulting in a more complex task switch than in Chapter 3. Potentially the between task effect size was small in both the behavioural and EEG data, as the low and high load task switch were equally demanding on available resources.

In combination, the results provide two important aspects of attention that impact the interpretation of the effect of a previous task demands on a continuous task requiring sustained attention. *1)* Attention-related alpha activity and frontal theta activity in combination with performance metrics, during a continuous task, are modulated by a previous task of differing attentional demand, and *2)* attention-related alpha manifests distinctly depending on the nature of the continuous task. In this thesis, I suggest that if the sustained visuomotor task is complex and demanding, then decrements in attention-related mechanisms manifest during the task. This is represented by reduced alpha ERD (i.e. not enough attentional resources). In addition, this should be coupled with reduced performance measures. However, if the secondary task is less demanding, then the attention resources required for successful task completion do not reach full attentional capacity, and therefore manifest as a greater increase in alpha ERD (i.e. enough attentional resources). This may not be paired with reduced performance.

Overall, these results provide insight into the attentional limitations at a neurobehavioural level during sustained tasks. The evidence suggests a neurophysiological response and cognitive limitations during multitasking scenarios, in particular task switching, which manifests differently depending on the nature of the primary and secondary task. As these studies emphasise that the direction of results are associated with the nature of the tasks, this provides motivation for future research measuring cognitive limitations in applied situations, which are more complex, demanding, and impacted by factors such as the internal state of a participant (e.g. motivation) and the environment.

### 7.2.2 Capturing and modulating lapses in sustained attention during semi-naturalistic and naturalistic studies of autonomous driving (Chapter 5 and 6)

#### *7.2.2.1 Overview*

Chapter 5 aimed to investigate the impact of an unexpected event during Level 5 autonomous driving on visual attention and autonomic arousal. Participants undertook several autonomous journeys that consisted of a stop: one unexpected stop initiated by a hazardous event, and one expected stop initiated by the journey setup. The results suggested that similar visual attention demands were expended but distinctly allocated during the unexpected and expected journeys. During an unexpected event, participants demonstrated increased sympathetic nervous activity and a smaller distribution of ocular fixations, coupled with longer fixation duration towards the hazard. These results imply a narrowing of focused attention during an adverse situation, which lasted once the critical event had ended. The study also demonstrated that the physiological indices examined can be considered useful and practical measures for evaluating functional state during real-world autonomous driving. A driver state monitoring (DSM) system, introduced in Chapter 1.3.1, could include such measures to detect these behaviours and make an informed decision on vehicle behaviour and adapt vehicle notifications accordingly.

The final study described in Chapter 6 aimed to investigate the attentional demands and user acceptability of a DSM notification during autonomous driving. The study was undertaken in a virtual autonomous driving simulator, which investigated ocular attention, physiological arousal, and user acceptability, during two types of notifications which preceded the vehicle slowing down: a DSM notification indicating changes in physiological state, and a conventional message of a speed limit change. Participants rated the reliability of the vehicle the highest after the DSM journey. Qualitative results suggested that participants found the notification useful and appropriate but sought control of the DSM system. There were no differences in the real-time physiological response between notifications, and eye fixation results revealed that the DSM notification was no more visually attentionally demanding compared to a speed limit notification. Although fixation duration and count were greater on the driving environment during and after a DSM notification compared to a speed limit change notification, this result was also significant before the notification. This suggests that it was not the notification that increased fixation duration. In combination, these results suggest that participants were not adversely affected nor visually distracted by the DSM notification, and generally accepted the systems capabilities. However, it was not clear whether the notification improved visual monitoring of the driving environment.

*7.2.2.2 Discussion*

There are two distinct differences between Chapter 5 and Chapter 6. While Chapter 5 explored older adults during real-world autonomous journeys, Chapter 6 explored younger adults during simulated autonomous journeys. Considering the potential autonomous vehicle benefits for older adults, such as maintaining mobility and independence, and the age-related individual differences related to human-automation collaboration, a comprehensive understanding of older adults' psychophysiological state during periods of automated driving, particularly during safety-critical situations, is needed. Previous research has indicated that during periods of automated driving, younger adults fixated on safety-critical areas, despite having no control of the vehicle (Strauch et al., 2019). Chapter 5 expands on this research by demonstrating that older adults also fixate on a safety-critical event, and this narrowed focus of attention lasts once the event had ended. This result is similar to Chapter 2, 3 and 4 that suggest that the demand of a previous task, e.g. an unexpected event in Chapter 5, impacts attention-related mechanisms during a continuous task.

Despite these disparities, similar self-reports for perceived situation awareness, trust, reliability, and workload before and after all journeys between both populations were found (see Table 7.2). For example, situation awareness ratings did not differ between pre- and post- journey ratings, in both younger adults in a simulator environment (pre: 19.55; post: 22.35), and older adults in a real-world environment (pre: 22.15; post: 21.41). These results contradict the literature that suggest older adults have lower situation awareness when compared to younger adults (e.g. Bolstad, 2001; Caserta & Abrams, 2007), and high levels of automation encourage driver complacency resulting in a decline in situation awareness (McCall et al., 2019). However, the non-significant results may have been driven by the situation awareness measurement technique – the self-report post-trial 'SART'. Salmon et al. (2009) suggests that if the task is dynamic and changeable, and the situation awareness content is not known, then the SART may not be the best method to study situation awareness. This is because the SART produces a score of how aware participants felt during the task. Therefore, accurate memory recall can impact the results. In addition, the situation awareness content is not known and so it is not clear what an ideal SART score is during autonomous driving. Therefore, it is difficult to compare a 'pre-' period, when no task had been undertaken, to a 'post-' period, when a task had been undertaken.. Trust and reliability both increased from pre- to post- journeys in both Chapters. Interestingly, trust and reliability ratings were generally lower in the younger adults (Trust: pre = 5, post = 8; Reliability: pre = 5, post = 7) when compared to the older adults (Trust and reliability: pre = 7, post = 9). Although some have suggested that the acceptance of using new technologies decreases with age (Kantowitz et al., 1993), a number of studies now suggest that older adults trust and over-rely on automation more than younger adults. For example, McBride et al. (2010) compared younger and older adults undertaking dual tasks with automation support. The automation support was incorrect 70% of the time, and participants were required to correct the automation. During automation errors, older adults were less likely to check the automation and self-report trust was greater, indicating potentially lower perceived situation awareness and higher dependence on the automation. The literature suggests that several factors may be involved. Because of age-related cognitive changes such as deficits in basic attention, memory, and learning (Zanto & Gazzaley, 2014), and differences in experiences with technology (Barnard et al., 2013), older adults may be less likely to detect errors and failures in the system, leading to a better perceived experience.

*Table 7.2.* Self-reports of situation awareness, trust, reliability, and workload factors, pre-journeys to post-journeys in Chapter 5 and Chapter 6.

| Parameter | Chapter 5<br>Field study with older adults | Chapter 6<br>Simulator study with younger adults |
|---|---|---|
| Situation awareness | – | – |
| Trust | ↑ | ↑ |
| Reliability | ↑ | ↑ |
| **NASA-TLX** | | |
| Mental | – | ↓ |
| Physical | ↓ | ↓ |
| Temporal | – | – |
| Effort | ↓ | ↓ |
| Performance | – | – |
| Frustration | ↑ | – |

*Key:* ↑ represents the parameter increased from pre- to post- journeys; ↓ represents the parameter decreased from pre- to post- journeys; – represents no change from pre- to post- journeys.

Although the findings from Chapter 5 and Chapter 6 indicate some age-related differences, it is important to recognise the potential influence of the different testing environments. In Chapter 5, a naturalistic study was undertaken, which employed a real autonomous vehicle in a pedestrianised environment. In Chapter 6, a virtual-based environment was employed. Strauch et al. (2019) recently explored passengers' trust and gaze behaviour during a highway drive in a manual and autonomous vehicle in the field, and in a manual and autonomous driving simulator. Trust ratings were similar between manual and autonomous driving in both the simulator and field experiment. Although trust ratings corresponded between the simulator and field experiment, ocular behaviours differed. While no effect on gaze behaviour was found in the simulator experiment, the field experiment revealed a greater number of fixations on safety-relevant features more so during autonomous compared to manual driving. The authors suggest that an underlying factor affected by the driving environment, such as interest or perceived risk, may play a role in manipulating gaze behaviour. Therefore, the different environments employed in Chapter 5 and 6 may have impacted the results. Despite the clear benefits of undertaking field studies, the naturalistic study undertaken in Chapter 5 was time-consuming, resource-heavy, and relatively inflexible. The study was limited as it had to adhere to strict safety protocols such as a safety driver, low speed (3 – 5 mph), and marshals surrounding the vehicle (further discussed in Chapter 5.5.5). The vehicle was continually under the risk of sensors/systems failing, as well as being adversely impacted by factors such as bad weather and heavily trafficked areas. Issues like these have also been well documented in the literature (e.g. Nascimento et al., 2019; Stadler et al., 2019). Consequently, virtual environments can be considered a practical tool for exploring behaviours and attitudes to Level 5 autonomous systems.

Overall, Chapter 5 and 6 were successful in the overall objectives of ascertaining feasibility of physiological measurements for development into a driver state monitoring system in applied contexts. In addition, the results from Chapter 5 and Chapter 6 provide insight into how visual attention and autonomic arousal manifest during unexpected situations, and whether a driver state monitoring system may have the potential to improve safety and wellbeing during Level 5 driving. The evidence suggests a narrowing of visual attention and increases in autonomic arousal during unexpected events during vehicle automation, and monitoring real-time continuous physiological signals may be an acceptable way of modulating visual attention and arousal, subsequently improving driver safety and wellbeing.

## 7.3 Challenges and future directions

### 7.3.1 Individual differences

Neurophysiological and physiological measures utilised within this thesis are all susceptible to individual differences. For example, 10% of people are considered EDA non-responders, meaning they do not respond electrodermally (Braithwaite et al., 2012). If any type of physiological monitoring system, including a DSM system, is unable to measure skin conductance from all individuals, then this will impact the accuracy and safety of the monitoring method. Other physiological inter-individual differences have also been well documented. Quer et al. (2020) studied the variation in heart rate over the course of two years and found that although mean heart rate was 65 bpm, inter-individual differences ranged between 40 – 109 BPM. Similarly, alpha oscillatory activity is typically defined between 8.5 – 12.5 Hz, with a mean frequency of 10 Hz (Klimesch, 2012); however, alpha frequency varies between individuals. Haegens et al. (2014) study of 51 participants undertaking a resting state recording, visual grating, and an n-back task, demonstrated that individual alpha peak frequency ranged between 7 – 14 Hz, with a standard deviation of 2.8 Hz. They also noted that some participants did not display an alpha peak. So, although within-subject designs help to control for inter-individual variation, using a narrow range of frequencies for analysis may not truly capture ‘alpha’ activity as manifested in the population.

Research traditionally focuses on inter-individual differences; however, it is recognised that intra-individual variability (i.e. a transient, within-participant change) will also impact physiological measurements. Quer et al. (2020) two-year study found that 20% of participants experienced at least 1 week in which their resting heart rate was modulated by 10 bpm or more. In addition, they found that intra-variability was more common in women and particularly those of a childbearing age. Other factors, such as experience, can result in intra-individual differences. For example, greater experience in driving tends to lead to a reduction in the duration of certain types of fixations (Chapman & Underwood, 1998). Therefore, any human state monitoring system would need to take into consideration the changes of ocular behaviour with experience and time as well as being flexible with regards to intra-individual variability with other measures including heart rate. Further research analysing the physiological and neural correlates of intra-individual differences in behaviour (i.e. a trial-by-trial analysis in Chapter 3 and 4) would reveal a more detailed and dynamic picture of attentional fluctuations associated with variables such as time-on-task and arousal.

As recommended by Laborde et al. (2017), the thesis attempted to control for confounding variables that could impact autonomic function. In Chapters 2, 3, 4 and 6, participants filled in a demographic questionnaire asking information on their smoking status (Ashare et al., 2012), previous night's sleep duration (Stein & Pu, 2012), recent physical activity (Stanley et al., 2013), caffeine consumption, including coffee, chocolate, tea, energy drinks, and headache medication (Zimmermann-Viehoff et al., 2016), alcohol consumption (Quintana et al., 2013), and age and gender (Umetani et al., 1998). Chapters 3 and 4 also baseline corrected the EEG data to remove any within-subject variation, while Chapters 5 and 6 standardised autonomic arousal data. However, there are countless variables that are difficult to measure that can influence physiological signals. For example, neural oscillatory measurements are influenced by genetic factors and EEG power shows moderate to high heritability (Smit et al., 2005). Therefore, group data will not pick up on details within the individual data, and the large inter-individual variability of responses hinders the direct comparability between individuals. Consequently, any human physiological measuring system must not apply the same parameters for all users; the system must be tailored towards a specific user.

Inter- and intra- subject variation has a significant impact on the interpretation of physiological signals. A human state monitoring system will be susceptible to false alarms and missed positives if it does not consider these differences. Multivariate methods, rather than group averaging approaches, may be more appropriate. However, differences still introduce biases in the data, most prominently heterogeneity, and research has already indicated that individual differences in physiological signals already poses challenges for machine learning algorithms (Biagetti et al., 2018). Consequently, any human physiological measuring system must be tailored towards a specific user. Further research is needed integrating multimodal measures and multivariate methods to classify functional state during applied studies, to understand which indicators can be reliably collected and analysed during applied contexts innate with possible artefacts.

### 7.3.2 Ethical considerations

Automated methods of real-time, unobtrusive monitoring are not without potential challenges and concerns for users and society. Monitoring physiological state during interactions with automation leads to new ethical issues. To begin to understand the ethical acceptability of these systems, it is important to consider the research on the ethics involved in physiological monitoring, automated systems, and vulnerable populations. This section will focus on several ethical concerns for public acceptance of human monitoring, which could hinder the successful implementation of human physiological monitoring in many applied contexts.

Research has indicated users are concerned about privacy invasion during personal health monitoring (Mihailidis et al., 2008). As discussed above in Chapter 7.3.1, a monitoring system must take into consideration the user's individual response, rather than utilising a one fits all framework. If the system exposes sensitive information, this could potentially heighten negative feelings of shame or embarrassment. For example, a high resting heart rate has been associated with a negative stigma of weight (Vartanian et al., 2018). Therefore, if a DSM system responds to a user's resting heart rate, the user may not want to share their vehicle with others. If a measure can only provide limited evidence of human state, then accessing this data is ethically unjustifiable. Consequently, it is important to consider the correct balance between a user's privacy and the effectiveness of the measurements, as well as how alerts will be managed and generated.

Physiological monitoring may encourage a threat to autonomy. Monitoring may intrude on the user's free choice and decision making (Kang et al., 2010). In addition, it may heighten other ethical anxieties such as feelings of deception, objectification and infantilisation (Cahill et al., 2018). For example, a user may feel as though they are being treated like a child if a DSM system advises them that they are fatigued and should pull over. The results from Chapter 6 provide some insight into these concerns. Participants expressed that they wanted more control over the DSM system. Some participants stated that they wanted the choice to increase or decrease the vehicle speed or turn the system off altogether. Therefore, having an override setting that allows users to change vehicle behaviour may minimise this ethical concern.

It is important that design considers the specific needs to mitigate enhanced ethical risks to older adults and vulnerable populations. Obtaining informed consent may be difficult in vulnerable populations or in populations with reduced decision-making capacity (Kang et al., 2010) and it may be appropriate to involve users and their families directly in the decision-making process. If monitoring of physiological parameters, such as in-home systems, are available to older adults rather than other populations, the issue of stigmatisation and infantisation may be greatly amplified (Cahill et al., 2018). For example, research has indicated that the use of a hearing aid or a walker may enhance physical autonomy but also emphasises decreased independence, and promotes stigmatisation (Parette & Scherer, 2004; Ruusuvuori et al., 2019). A human state monitoring system, only available to older adults, may encourage similar adverse feelings. Another aspect to consider is the user's ability to utilise complex autonomous systems. Users may blame themselves for any difficulty they experience using the technology. These issues are potentially amplified in an older adult population with ageing-related impairments, as they are more likely to rely on automated systems (McBride et al., 2011), and are more prone to lack understanding of advanced technology (Mann et al., 2007). Developing technology with a better understanding of the needs and user preferences of older adults and other populations may resolve some ethical concerns.

The propensity for technology to rapidly advance and overtake ethical solutions is well recognised. Many ethical concerns including privacy, autonomy, and stigmatisation are associated with physiological monitoring and the use of complex systems. Although the majority of research aims to understand whether it is possible to monitor physiology in applied contexts, it is equally vital to understand how, when, and under which conditions human physiological monitoring should be used. Not considering such ethical concerns regarding how, when, and under which conditions, could hinder the implementation and public acceptance of human physiological monitoring in many applied contexts, including driver state monitoring.

### 7.3.3 The state of the technology

The wearable Empatica E4, previously used in the literature (e.g. Bitkina et al., 2019; Menghini et al., 2019; Walker et al., 2019), was used to collect signals of heart rate and electrodermal activity (EDA) in the applied studies described in Chapter 5 and Chapter 6. Despite the E4s success in detecting EDA differences between the unexpected stop in Chapter 5, the device does not come without its limitations. Conventional physiological research uses the distal or intermediate phalanges of the ring and index fingers (Boucsein et al., 2012). The E4, like many wearables, measures skin conductance via wrist sensors, and the issues with this are two-fold. Firstly, the wrist has fewer sweat glands than the fingers (Boucsein et al., 2012). In addition, sweating at the wrist is more likely during an affective response (Wilke et al., 2007). Therefore, the results of applied research may differ compared to laboratory-based work, and the signal will not have a high cross-correlation due to placement differences. As the wrist is less responsive to EDA, an underestimation of EDA parameters is expected (Payne et al., 2016; van Dooren et al., 2012). This may have led to the non-significant findings in Chapter 6. As the signal is smaller at the wrist, it is difficult to detect skin conductance responses following conventional parameters (e.g. 0.03 µs changes from the background signal; Braithwaite et al., 2013). Therefore, only skin conductance level was analysed during Chapters 5 and 6, but this would have also included the skin conductance response signal.

In addition, heart rate measures were collected by photoplethysmography (PPG) at the wrist, whereas conventional heart rate measures are collected via electrocardiography (ECG) via the chest. While ECG sensors measures the bio-potential generated by electrical signals that control the expansion and contraction of heart chambers; PPG use a light-based technology to sense the rate of blood flow as controlled by the heart's pumping action. Although heart rate is similar between the wrist and chest, the PPG signal is less accurate than ECG (Laborde et al., 2017). Shaffer et al. (2014) demonstrated that changes in pulse transit time from changes in the elasticity of the arteries cannot be captured by interbeat intervals. As such, a PPG signal will not reflect elasticity changes in the arteries, which have been associated with high arousal situations such as stress. In addition, interbeat interval data often suffers from missing samples, particularly during movement (Baek & Shin, 2017; Gambi et al., 2017). When samples are missing, successive differences cannot be calculated, which unfortunately was a challenge faced during Chapters 5 and 6. Moreover, Gruden et al. (2019) found a relative lack of sensitivity with the E4 which was affected by manual driving related movement artefacts, and it was unable to distinguish an easy driving scenario from a demanding driving scenario compared to a 3-channel ECG monitor based on standard HRV parameters. In Chapters 5 and 6, movement artefacts inherent to wearable devices were limited by asking participants to wear it on their non-dominant wrist while in a seated setting. As testing was in an autonomous driving environment, participants did not need to move their hands like in manual driving. Motion artefacts between conditions were analysed to ensure that there was not a significant difference between movement (see Appendix 5.4

and Appendix 6.2). Accelerometer values did differ between journeys during Chapter 6 potentially enlarging skin conductance and heart rate values. There was not a significant difference in physiology between journeys, and so it is possible that movement artefacts obstructed any differences. Despite this, the significant difference in movement demonstrates the importance of analysing accelerometer values when interpreting physiological data from wearable devices. Lastly, the Empatica E4 has a low sampling rate of 64 Hz. As poor sampling rates can distort waveform characteristics by inducing artificial oscillatory characteristics not part of the true signal, it is often cited that a sampling rate of at least 250 Hz is required for heart rate variability analysis (see Shaffer & Ginsberg, 2017; Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology, 1996), although recent research has indicated that 100 Hz is appropriate for time-domain heart rate variability features (Kwon et al., 2018). Nevertheless, a sampling rate of 64 Hz can be considered too low for accurate heart rate variability capture.

Overall, this thesis in combination with recent research, suggests that the Empatica E4 is not usable for sudden and short-lived stressors (van Lier et al., 2020). This is problematic for a physiological monitoring system which would need to detect small effects to sudden events. Further research is required to ensure non-contact, unobtrusive, and acceptable physiological recording devices can collect accurate and reliable devices in dynamic scenarios which are inherently full of sudden events, motion artefacts, and poor lighting conditions.

### 7.3.4 Big data analytics

A human state monitoring system must collect and integrate a set of behavioural and physiological indices simultaneously recorded in real time. Adding signals from multimodal sources, with potentially millisecond temporal resolution, will result in petabytes of data needing to be efficiently generated, processed, and interpreted. The fusion of data studying complex human activities provides big challenges for big data analytics. As discussed in Chapter 7.3.2, physiological monitoring comes with ethical issues related to confidentiality, security, and over-collection, which need to be considered for all streams of data. In addition, the variety of data formats, e.g. video data to infer facial expression, sensor data to infer human electrophysiological signals, comes with significant challenges related to their dissimilar form (Sivarajah et al., 2017). The issues surrounding individual differences are associated to all forms of human monitoring, and the combination of many different signals will result in a high rate of non-homogeneous data. In addition, Sivarajah et al. (2017) demonstrated that data in which meaning rapidly changes also proves a challenge for big data analytics. For example, data should be interpreted differently when heart rate increases during traffic, and when heart rate increases coming over a brow of a hill. In these instances, it would not be appropriate for a DSM system to change vehicle behaviour on the physiological signals alone. The challenges facing big data analytics in human physiological monitoring may negate any benefits of a monitoring system. Therefore, further research is needed to uncover how to manage and interpret large, varied, complex, non-homogenous datasets, before successful monitoring approaches can be implemented.

### 7.3.5 Other future directions

While the above challenges suggest important research avenues for future research, the overall findings from this thesis also offers suggestions for further research. While the initial three experimental Chapters provoke questions regarding the mechanisms involved in performance deficits during sustained attention, the last two Chapters provoke questions regarding the implementation of monitoring performance deficits during applied sustained attention tasks. Some further research questions and future directions are discussed below.

The General Introduction Chapter revealed the potential for real-time neurofeedback to influence underlying neural networks to improve attentional deficits in clinical populations (Chapter 1.3.2). Chapters 2, 3, and 4 findings revealed a vital role of a fronto-parietal network involved in sustained attention, and in particular, a vital role of parieto-occipital alpha activity in task performance, demonstrating lower alpha values (increased alpha ERD) with better performance (Chapter 4). As lasting sustained attention deficits are pervasive in a broad range of clinical populations, further research should determine whether combined EEG-fMRI neurofeedback could be used to self-regulate alpha activity in the parieto-occipital cortices in patient populations. Using a combined approach will improve the specificity, so self-regulation can focus on relevant oscillatory activity in the relevant structures to improve performance deficits during sustained attention. In addition, as Chapter 4 revealed a continuous impact on parieto-occipital alpha activity following a task switch, it would be interesting to see whether neurofeedback in a healthy cohort could increase alpha ERD following a task switch, and whether this has any impact on task performance.

Furthermore, transcranial magnetic stimulation studies are needed to infer the regions and connections that are essential to the maintenance of sustained attention. Changes in task performance that result from stimulation of brain activity can provide insight into the causal role of that activity, whereas other imaging techniques, such as EEG, can only provide correlational information about the relationship between behaviour and brain activity. Short trains of repetitive transcranial magnetic stimulation in the alpha frequency band has been shown to improve task performance and task-related alpha desynchronisation (Klimesch et al., 2003). Further research could determine whether repetitive transcranial magnetic stimulation in the alpha frequency band impacts continuous performance measures during more naturalistic tasks such as simulated driving, providing causal evidence for a role of alpha oscillations during naturalistic tasks.

The last two experimental Chapters focused of the feasibility of a biological-based driver state monitoring system during autonomous driving. Yet, the literature indicates that the driver's state also depends on various interrelated factors such as the external driving environment and the level of automation, for example, traffic density and weather have been shown to impact takeover performance (Gold et al., 2016; Li et al., 2018). Further research is needed to explore the relation between these factors and driver state. For example, if a driver demonstrates fatigued behaviours (i.e. an increase in eye blinks) during adverse weather conditions, would driving performance be impaired when compared to clear weather conditions? This is an important avenue of research to further inform a driver assist system adapting vehicle behaviour and notifications.

### 7.4 Conclusion

This thesis set out to explore the potential role of physiological indices as adjunctive measures of continuous performance and sustained attention. This was with a view for future developments of driver monitoring systems to improve transport safety and wellbeing. In addition, this thesis aimed to explore how the engagement of a previous task could impact sustained attention and task performance and processes, and whether fluctuations could be captured reliably during naturalistic settings. As decrements in sustained attention are multifaceted and operate on a series of timescales, physiological and neurophysiological indices were used to capture fluctuations in attention during naturalistic settings. Despite the promises of a real-time continuous measurement system, **Chapter 2** revealed that subjective ratings were the only sensitive measure of attentional load during multitasking. However, the experiment also revealed a carry-over neural effect of multitasking, in the form of increased frontal theta power, likely reflecting ongoing cognitive processes following task cessation. Moreover, **Chapter 3** demonstrated a greater suppression of parietal alpha activity during a continuous tracking task following a task of high attentional load, likely reflecting a stronger recruitment of visuomotor attentional processes. In addition, accuracy was negatively impacted following a task of high load. On the contrary, **Chapter 4** demonstrated reduced suppression of parieto-occipital alpha activity following a task of high load when compared to a passive viewing task, likely reflecting deficits in attentional resource allocation. Furthermore, this was supported by neurobehavioral correlations associating lower alpha power with increased accuracy. In conclusion, **Chapters 3 and 4** suggest that performance decrements during a complex visuomotor task such as driving, following an unrelated task of varying attentional load, manifest as reduced parieto-occipital alpha desynchronisation (i.e. greater alpha power). As such, these studies were successful in providing evidence for the impact of prior attentional load on sustained attention and continuous performance. Narrowed attention during a naturalistic sustained attention task, autonomous driving, was revealed in **Chapter 5**, demonstrated by an increased number and duration of fixations, and sympathetic dominance, during an unexpected event. Furthermore, the findings from **Chapter 6** suggests continual monitoring of autonomic arousal may be an acceptable and unobtrusive approach to modulate behaviour and wellbeing, and notifications of driver state are no more visually demanding than driving-related notifications. However, it is not clear whether feedback of driver state can successfully modulate visual attention, and so further research is needed. Overall, the findings of this thesis are of interest to inform future research to facilitate improved understanding of attentional deficits in healthy and clinical populations. In addition, this thesis supports future research to facilitate development of technologies such as brain-computer interfaces to modulate time-varying physiological processes to improve safety and wellbeing during sustained attention tasks in everyday and occupational settings.

# References


Affanni, A., Bernardini, R., Piras, A., Rinaldo, R., & Zontone, P. (2018). Driver's stress detection using Skin Potential Response signals. *Measurement: Journal of the International Measurement Confederation*, *122*, 264–274. https://doi.org/10.1016/j.measurement.2018.03.040

Akash, K., Hu, W. L., Jain, N., & Reid, T. (2018). A classification model for sensing human trust in machines using EEG and GSR. *ACM Transactions on Interactive Intelligent Systems*, *8*(4), 1–20. https://doi.org/10.1145/3132743

Al-Hashimi, O., Zanto, T. P., & Gazzaley, A. (2015). Neural sources of performance decline during continuous multitasking. *Cortex*, *71*, 49–57. https://doi.org/10.1016/j.cortex.2015.06.001

Alberdi, A., Aztiria, A., & Basarab, A. (2016). Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review. *Journal of Biomedical Informatics*, *59*, 49–75. https://doi.org/10.1016/j.jbi.2015.11.007

Allport, A., Styles, E., & Hsieh, S. (1994). Shifting Intentional Set: Exploring the Dynamic Control of Tasks. In *C. Umilta & M. Moscovitch (Eds.), Attention and Performance XV*. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/1478.003.0025

Alrefaie, M. T., Summerskill, S., & Jackon, T. W. (2019). In a heart beat: Using driver's physiological changes to determine the quality of a takeover in highly automated vehicles. *Accident Analysis and Prevention*, *131*, 180–190. https://doi.org/10.1016/j.aap.2019.06.011

Anderson, J. R., Bothell, D., & Douglass, S. (2004). Eye Movements Do Not Reflect Retrieval Processes: Limits of the Eye-Mind Hypothesis. *Psychological Science*, *15*(4), 225–231. https://doi.org/10.1111/j.0956-7976.2004.00656.x

Anderson, K. L., & Ding, M. (2011). Attentional modulation of the somatosensory mu rhythm. *Neuroscience*, *180*, 165–180. https://doi.org/10.1016/j.neuroscience.2011.02.004

Andreassi, J. L. (2006). Electrodermal activity and behavior. In *J. L. Andreassi (Ed.), Psychophysiology: Human Behavior and Physiological Response* (pp. 191–218). Mahwah, NJ: Lawrence Erlbaum.

Anstey, K. J., Windsor, T. D., Luszcz, M. A., & Andrews, G. R. (2006). Predicting driving cessation over 5 years in older adults: Psychological well-being and cognitive competence are stronger predictors than physical health. *Journal of the American Geriatrics Society*, *54*(1), 121–126. https://doi.org/10.1111/j.1532-5415.2005.00471.x

Arexis, M., Gaspelin, N., & Ruthruff, E. (2017). Attentional Capture in Driving Displays. *British Journal of Psychology*, *108*(2), 259–275. https://doi.org/10.1111/bjop.12197

Arizpe, J., Kravitz, D. J., Yovel, G., & Baker, C. I. (2012). Start position strongly influences fixation patterns during face processing: Difficulties with eye movements as a measure of information use. *PLoS ONE*, *7*(2), e31106. https://doi.org/10.1371/journal.pone.0031106

Ashare, R. L., Sinha, R., Lampert, R., Weinberger, A. H., Anderson, G. M., Lavery, M. E., Yanagisawa, K., & McKee, S. A. (2012). Blunted vagal reactivity predicts stress-precipitated tobacco smoking. *Psychopharmacology*, *220*(2), 259–268. https://doi.org/10.1007/s00213-011-2473-3

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Athènes, S., Chatty, S., Bustico, A., & Bustico, R. (2000). Human factors in ATC alarms and notifications design: an experimental evaluation. *Proc. ATM'2000 R&D Seminar*, *3*, 1–6. https://doi.org/10.1.1.35.7529

Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*(1), 201–208. https://doi.org/10.1016/j.neuroscience.2005.08.023

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Babiloni, C., Carducci, F., Cincotti, F., Rossini, P. M., Neuper, C., Pfurtscheller, G., & Babiloni, F. (1999). Human movement-related potentials vs desynchronization of EEG alpha rhythm: A high-resolution EEG study. *NeuroImage*, *10*(6), 658–665. https://doi.org/10.1006/nimg.1999.0504

Baek, H. J., & Shin, J. W. (2017). Effect of Missing Inter-Beat Interval Data on Heart Rate Variability Analysis Using Wrist-Worn Wearables. *Journal of Medical Systems*, *41*(10), 147. https://doi.org/10.1007/s10916-017-0796-2

Bailey, L. M., McMillan, L. E., & Newman, A. J. (2020). A sinister subject: Quantifying handedness‐based recruitment biases in current neuroimaging research. *European Journal of Neuroscience, 51*(7), 1642-1656

Baldwin, C. L., Roberts, D. M., Barragan, D., Lee, J. D., Lerner, N., & Higgins, J. S. (2017). Detecting and quantifying mind wandering during simulated driving. *Frontiers in Human Neuroscience*, *11*, 406. https://doi.org/10.3389/fnhum.2017.00406

Barceló, F., & Cooper, P. S. (2018). Quantifying contextual information for cognitive control. *Frontiers in Psychology*, *9*, 1–4. https://doi.org/10.3389/fpsyg.2018.01693

Barnard, M. P., & Chapman, P. (2018). The effects of instruction and environmental demand on state anxiety, driving performance and autonomic activity: Are ego-threatening manipulations effective? *Transportation Research Part F: Traffic Psychology and Behaviour*, *55*, 123–135. https://doi.org/10.1016/j.trf.2018.02.040

Barnard, Y., Bradley, M. D., Hodgson, F., & Lloyd, A. D. (2013). Learning to use new technologies by older adults: Perceived difficulties, experimentation behaviour and usability. *Computers in Human Behavior*, *29*(4), 1715–1724. https://doi.org/10.1016/j.chb.2013.02.006

Bauer, M., Oostenveld, R., Peeters, M., & Fries, P. (2006). Tactile spatial attention enhances gamma-band activity in somatosensory cortex and reduces low-frequency activity in parieto-occipital areas. *Journal of Neuroscience*, *26*(2), 490–501. https://doi.org/10.1523/JNEUROSCI.5228-04.2006

Beauchamp, M. S., Petit, L., Ellmore, T. M., Ingeholm, J., & Haxby, J. V. (2001). A parametric fMRI study of overt and covert shifts of visuospatial attention. *NeuroImage*, *14*(2), 310–321. https://doi.org/10.1006/nimg.2001.0788

Bédard, M., Parkkari, M., Weaver, B., Riendeau, J., & Dahlquist, M. (2010). Assessment of driving performance using a simulator protocol: Validity and reproducibility. *American Journal of Occupational Therapy*, *64*(2), 336–340. https://doi.org/10.5014/ajot.64.2.336

Behrmann, M., Geng, J. J., & Shomstein, S. (2004). Parietal cortex and attention. *Current Opinion in Neurobiology*, *14*(2), 212–217. https://doi.org/10.1016/j.conb.2004.03.012

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Benoni, H., & Tsal, Y. (2010). Where have we gone wrong? Perceptual load does not affect selective attention. *Vision Research*, *50*(13), 1292–1298. https://doi.org/10.1016/j.visres.2010.04.018

Betz, T., Kietzmann, T. C., Wilming, N., & König, P. (2010). Investigating task-dependent top-down effects on overt visual attention. *Journal of Vision*, *10*(3), 1–14. https://doi.org/10.1167/10.3.15

Biagetti, G., Crippa, P., Falaschetti, L., Tanoni, G., & Turchetti, C. (2018). A comparative study of machine learning algorithms for physiological signal classification. *Procedia Computer Science*, *26*, 1977–1984. https://doi.org/10.1016/j.procS.2018.07.255

Bishop, S. J., Jenkins, R., & Lawrence, A. D. (2007). Neural processing of fearful faces: Effects of anxiety are gated by perceptual capacity limitations. *Cerebral Cortex*, *17*(7), 1595–1603. https://doi.org/10.1093/cercor/bhl070

Bitkina, O. V., Kim, J., Park, J., Park, J., & Kim, H. K. (2019). Identifying traffic context using driving stress: A longitudinal preliminary case study. *Sensors*, *19*(9), 2152. https://doi.org/10.3390/s19092152

Bogadhi, A. R., Bollimunta, A., Leopold, D. A., & Krauzlis, R. J. (2019). Spatial Attention Deficits Are Causally Linked to an Area in Macaque Temporal Cortex. *Current Biology*, *29*(5), 726–736. https://doi.org/10.1016/j.cub.2019.01.028

Bologva, E. V., Prokusheva, D. I., Krikunov, A. V., Zvartau, N. E., & Kovalchuk, S. V. (2016). Human-Computer Interaction in Electronic Medical Records: From the Perspectives of Physicians and Data Scientists. *Procedia Computer Science*, *100*, 915–920. https://doi.org/10.1016/j.procs.2016.09.248

Bolstad, C. A. (2001). Situation awareness: Does it change with age? *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, 272–276. https://doi.org/10.1177/154193120104500401

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

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

Boutcher, Y. N., & Boutcher, S. H. (2006). Cardiovascular response to Stroop: Effect of verbal response and task difficulty. *Biological Psychology*, *73*(3), 235–241. https://doi.org/10.1016/j.biopsycho.2006.04.005

Bradley, M. M., & Lang, P. J. (2002). Measuring Emotion: Behavior, Feeling, and Physiology. In *R. D. Lane & L. Nadel (Eds.), Cognitive neuroscience of emotion* (pp. 49–59). New York, NY: Oxford University Press.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Braithwaite, J. J., Broglia, E., & Watson, D. G. (2014). Autonomic emotional responses to the induction of the rubber-hand illusion in those that report anomalous bodily experiences: Evidence for specific psychophysiological components associated with illusory body representations. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1131–1145. https://doi.org/10.1037/a0036077

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe Biopac, M. (2012). A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments. *Psychophysiology*, *49*(1), 1017–1034. https://www.birmingham.ac.uk/Documents/college-les/psych/saal/guide-electrodermal-activity.pdf

Brefczynski-Lewis, J. A., Lutz, A., Schaefer, H. S., Levinson, D. B., & Davidson, R. J. (2007). Neural correlates of attentional expertise in long-term meditation practitioners. *Proceedings of the*

*National Academy of Sciences of the United States of America*, *104*(27), 11483–11488. https://doi.org/10.1073/pnas.0606552104

Broadbent, D. E. (1958). *Perception and communication*. New York, NY: Pergamon Press. https://doi.org/10.1108/eb015727

Broeker, L., Haeger, M., Bock, O., Kretschmann, B., Ewolds, H., Künzell, S., & Raab, M. (2020). How visual information influences dual-task driving and tracking. *Experimental Brain Research*, *238*(3), 675–687. https://doi.org/10.1007/s00221-020-05744-8

Brooks, J. R., Garcia, J. O., Kerick, S. E., & Vettel, J. M. (2016). Differential functionality of right and left parietal activity in controlling a motor vehicle. *Frontiers in Systems Neuroscience*, *10*, 1–11. https://doi.org/10.3389/fnsys.2016.00106

Broyd, S. J., Demanuele, C., Debener, S., Helps, S. K., James, C. J., & Sonuga-Barke, E. J. S. (2009). Default-mode brain dysfunction in mental disorders: A systematic review. *Neuroscience and Biobehavioral Reviews*, *33*(3), 279–296. https://doi.org/10.1016/j.neubiorev.2008.09.002

Brunyé, T. T., & Gardony, A. L. (2017). Eye tracking measures of uncertainty during perceptual decision making. *International Journal of Psychophysiology*, *120*, 60–68. https://doi.org/10.1016/j.ijpsycho.2017.07.008

Brzezicka, A., Kamiński, J., Reed, C. M., Chung, J. M., Mamelak, A. N., & Rutishauser, U. (2018). Working memory load-related theta power decreases in dorsolateral prefrontal cortex predict individual differences in performance. *Journal of Cognitive Neuroscience*, *31*(9), 1290–1307. https://doi.org/10.1162/jocn_a_01417

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1–38. https://doi.org/10.1196/annals.1440.011

Bueno, M., Dogan, E., Selem, F. H., Monacelli, E., Boverie, S., & Guillaume, A. (2016). How different mental workload levels affect the take-over control after automated driving. *IEEE 19th International Conference on Intelligent Transportation Systems*, 2040–2045. https://doi.org/10.1109/ITSC.2016.7795886

Cahill, J., McLoughlin, S., & Wetherall, S. (2018). The Design of New Technology Supporting Wellbeing, Independence and Social Participation, for Older Adults Domiciled in Residential Homes and/or Assisted Living Communities. *Technologies*, *6*(1), 18. https://doi.org/10.3390/technologies6010018

Campagne, A., Pebayle, T., & Muzet, A. (2004). Correlation between driving errors and vigilance level: Influence of the driver's age. *Physiology and Behavior*, *80*(4), 515–524. https://doi.org/10.1016/j.physbeh.2003.10.004

Canino-Rodríguez, J. M., García-Herrero, J., Besada-Portas, J., Ravelo-García, A. G., Travieso-González, C., & Alonso-Hernández, J. B. (2015). Human computer interactions in next-generation of aircraft smart navigation management systems: Task analysis and architecture under an agent-oriented methodological approach. *Sensors*, *15*(3), 5228–5250. https://doi.org/10.3390/s150305228

Capotosto, P., Babiloni, C., Romani, G. L., & Corbetta, M. (2009). Frontoparietal cortex controls spatial attention through modulation of anticipatory alpha rhythms. *Journal of Neuroscience*, *29*(18), 5863–5872. https://doi.org/10.1523/JNEUROSCI.0539-09.2009

Carrier, L. M., Rosen, L. D., Cheever, N. A., & Lim, A. F. (2015). Causes, effects, and practicalities of everyday multitasking. *Developmental Review*, *35*, 64–78. https://doi.org/10.1016/j.dr.2014.12.005

Carsten, O., Lai, F., Barnard, Y., Jamson, H., & Merat, N. (2012). Control task substitution in semiautomated driving: does it matter what aspects are automated? *Human Factors*, *54*(5), 747–761.

Cascio, W. F., & Montealegre, R. (2016). How Technology Is Changing Work and Organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, *3*, 349–375. https://doi.org/10.1146/annurev-orgpsych-041015-062352

Caserta, R. J., & Abrams, L. (2007). The relevance of situation awareness in older adults' cognitive functioning: A review. *European Review of Aging and Physical Activity*, *4*(1), 3–13. https://doi.org/10.1007/s11556-007-0018-x

Castiello, U., & Paine, M. (2002). Effects of left parietal injury on covert orienting of attention. *Journal of Neurology Neurosurgery and Psychiatry*, *72*, 73–76. https://doi.org/10.1136/jnnp.72.1.73

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

Chajut, E., & Algom, D. (2003). Selective Attention Improves Under Stress: Implications for Theories of Social Cognition. *Journal of Personality and Social Psychology*, *85*(2), 231–248. https://doi.org/10.1037/0022-3514.85.2.231

Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception*, *27*(8), 951–964. https://doi.org/10.1068/p270951

Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, *363*, 345–347. https://doi.org/10.1038/363345a0

Chen, W., Sawaragi, T., & Horiguchi, Y. (2019). Measurement of Driver's Mental Workload in Partial Autonomous Driving. *IFAC-PapersOnLine*, *52*(19), 347–352.

https://doi.org/10.1016/j.ifacol.2019.12.083

Cheyne, J, A., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*, *111*(1), 98–113. https://doi.org/10.1016/j.cognition.2008.12.009

Chien, Y. L., Gau, S. S. F., Chiu, Y. N., Tsai, W. C., Shang, C. Y., & Wu, Y. Y. (2014). Impaired sustained attention, focused attention, and vigilance in youths with autistic disorder and Asperger's disorder. *Research in Autism Spectrum Disorders*, *8*(7), 881–889. https://doi.org/10.1016/j.rasd.2014.04.006

Choi, J. K., & Ji, Y. G. (2015). Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *International Journal of Human-Computer Interaction*, *31*(10), 692–702. https://doi.org/10.1080/10447318.2015.1070549

Chuang, C. H., Cao, Z., King, J. T., Wu, B. S., Wang, Y. K., & Lin, C. T. (2018). Brain electrodynamic and hemodynamic signatures against fatigue during driving. *Frontiers in Neuroscience*, *12*, 1–12. https://doi.org/10.3389/fnins.2018.00181

Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia*, *49*(6), 1407–1409. https://doi.org/10.1016/j.neuropsychologia.2011.01.029

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. (1988). *Statistical power analysis for the behavioural sciences*. Hillside, NJ: Lawrence Earlbaum Associates. https://doi.org/10.1111/1467-8721.ep10768783

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 115–159. https://doi.org/10.1037/0033-2909.112.1.155

Cohen, M. (2014). *Analyzing Neural Time Series Data*. Cambridge, MA: MIT Press. http://books.google.com.au/books?id=jTSkAgAAQBAJ&pg=PA132&dq=intitle:Analyzing+Neural+Time+Series+Data+Theory+and+Practice&hl=&cd=1&source=gbs_api%5Cnpapers2://publication/uuid/D12F703F-8A9A-4E57-A8DB-49571FE0DDB5%5Cnhttps://mitpress.mit.edu/books/analyzi

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*(131), 1–12. https://doi.org/10.3389/fnhum.2019.00131

Comte, S. L. (2000). New systems: New behaviour? *Transportation Research Part F: Traffic Psychology and Behaviour*, *3*(2), 95–111. https://doi.org/10.1016/S1369-8478(00)00019-X

Connor, J., Norton, R., Ameratunga, S., Robinson, E., Civil, I., Dunn, R., Bailey, J., & Jackson, R. (2002).

Driver sleepiness and risk of serious injury to car occupants: Population based case control study. *British Medical Journal*, *324*(7346), 1125. https://doi.org/10.1136/bmj.324.7346.1125

Cooper, P. S., Darriba, Á., Karayanidis, F., & Barceló, F. (2016). Contextually sensitive power changes across multiple frequency bands underpin cognitive control. *NeuroImage*, *132*, 499–511. https://doi.org/10.1016/j.neuroimage.2016.03.010

Cooper, P. S., Karayanidis, F., McKewen, M., McLellan-Hall, S., Wong, A. S. W., Skippen, P., & Cavanagh, J. F. (2019). Frontal theta predicts specific cognitive control-induced behavioural changes beyond general reaction time slowing. *NeuroImage*, *189*, 130–140. https://doi.org/10.1016/j.neuroimage.2019.01.022

Cooper, P. S., Wong, A. S. W., McKewen, M., Michie, P. T., & Karayanidis, F. (2017). Frontoparietal theta oscillations during proactive control are associated with goal-updating and reduced behavioral variability. *Biological Psychology*, *129*, 253–264. https://doi.org/10.1016/j.biopsycho.2017.09.008

Corbetta, M. (1998). Frontoparietal cortical networks for directing attention and the eye to visual locations: Identical, independent, or overlapping neural systems? *Proceedings of the National Academy of Sciences of the United States of America*, *95*(3), 831–838. https://doi.org/10.1073/pnas.95.3.831

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*, *58*(3), 306–324. https://doi.org/10.1016/B978-012370877-9.00163-8

Correa, Á., Molina, E., & Sanabria, D. (2014). Effects of chronotype and time of day on the vigilance decrement during simulated driving. *Accident Analysis and Prevention*, *67*, 113–118. https://doi.org/10.1016/j.aap.2014.02.020

Criswell, E. (2011). *Cram's Introduction to Surface Electromyography*. Sudbury, MA: Jones & Bartlett Publishers. https://doi.org/10.1002/car.1158

Croft, R. J., & Barry, R. J. (2000). Removal of ocular artifact from the EEG: A review. *Neurophysiologie Clinique*, *30*(1), 5–19. https://doi.org/10.1016/S0987-7053(00)00055-1

Cunillera, T., Fuentemilla, L., Periañez, J., Marco-Pallarès, J., Krämer, U. M., Càmara, E., Münte, T. F., & Antoni, R. F. (2012). Brain oscillatory activity associated with task switching and feedback processing. *Cognitive, Affective and Behavioral Neuroscience*, *12*(1), 16–33. https://doi.org/10.3758/s13415-011-0075-5

Davidson, M. A. (2008). ADHD in adults: A review of the literature. *Journal of Attention Disorders*, *11*(6), 628–641. https://doi.org/10.1177/1087054707310878

Dawson, M., Schell, A. M., & Filion, D. L. (2007). The Electrodermal System. In *J. T. Cacioppo, L. G.*

Tassinary & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (pp. 159–181). New York, NY: Cambridge University Press. https://doi.org/http://dx.doi.org/10.1017/CBO9780511546396.007

de Haan, B., Morgan, P. S., & Rorden, C. (2008). Covert orienting of attention and overt eye movements activate identical brain regions. *Brain Research*, *1204*, 102–111. https://doi.org/10.1016/j.brainres.2008.01.105

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

Deeney, C., & O'Sullivan, L. W. (2017). Effects of cognitive loading and force on upper trapezius fatigue. *Occupational Medicine*, *67*(9), 678–683. https://doi.org/10.1093/occmed/kqx157

DeGutis, J., Grosso, M., & VanVleet, T. (2017). Sustained Attention Training Reduces Spatial Bias in Parkinson's Disease: A Pilot Case Series. *Physiology & Behavior*, *176*(12), 139–148. https://doi.org/10.1016/j.physbeh.2017.03.040

Dellinger, A. M., Sehgal, M., Sleet, D. A., & Barrett-Connor, E. (2001). Driving cessation: What older former drivers tell us. *Journal of the American Geriatrics Society*, *49*(4), 431–435. https://doi.org/10.1046/j.1532-5415.2001.49087.x

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some Theoretical Considerations. *Psychological Review*, *70*(1), 80–90. https://doi.org/10.1037/h0039515

Dickerson, S. S., & Kemeny, M. E. (2004). Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychological Bulletin*, *130*(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Dickmeis, T. (2009). Glucocorticoids and the circadian clock. *Journal of Endocrinology*, *200*(1), 3–22. https://doi.org/10.1677/JOE-08-0415

Dimitrijevic, A., Smith, M. L., Kadis, D. S., & Moore, D. R. (2017). Cortical alpha oscillations predict speech intelligibility. *Frontiers in Human Neuroscience*, *11*(88), 1–10. https://doi.org/10.3389/fnhum.2017.00088

Doesburg, S. M., Bedo, N., & Ward, L. M. (2016). Top-down alpha oscillatory network interactions during visuospatial attention orienting. *NeuroImage*, *132*, 512–519. https://doi.org/10.1016/j.neuroimage.2016.02.076

Dogan, E., Honnêt, V., Masfrand, S., & Guillaume, A. (2019). Effects of non-driving-related tasks on takeover performance in different takeover situations in conditionally automated driving. *Transportation Research Part F: Psychology and Behaviour*, *62*, 494–504. https://doi.org/10.1016/j.trf.2019.02.010

Dong, Y., Hu, Z., Uchimura, K., & Murayama, N. (2011). Driver inattention monitoring system for intelligent vehicles: A review. *IEEE Transactions on Intelligent Transportation Systems*, *12*(2), 596–614. https://doi.org/10.1109/TITS.2010.2092770

Doricchi, F., MacCi, E., Silvetti, M., & MacAluso, E. (2010). Neural correlates of the spatial and expectancy components of endogenous and stimulus-driven orienting of attention in the posner task. *Cerebral Cortex*, *20*(7), 1574–1585. https://doi.org/10.1093/cercor/bhp215

Dosenbach, N. U. F., Fair, D. A., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences*, *12*(3), 99–105. https://doi.org/10.1016/j.tics.2008.01.001

Dosenbach, N. U. F., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A. T., Fox, M. D., Snyder, A. Z., Vincent, J. L., Raichle, M. E., Schlaggar, B. L., & Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(26), 11073–11078. https://doi.org/10.1073/pnas.0704320104

Dosenbach, N. U. F., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., Burgund, E. D., Grimes, A. L., Schlaggar, B. L., & Petersen, S. E. (2006). A Core System for the Implementation of Task Sets. *Neuron*, *50*(5), 799–812. https://doi.org/10.1016/j.neuron.2006.04.031

Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & Yves Von Cramon, D. (2000). Prefrontal cortex activation in task switching: An event-related fMRI study. *Cognitive Brain Research*, *9*(1), 103–109. https://doi.org/10.1016/S0926-6410(99)00029-4

Downey, L., & Van Willigen, M. (2005). Environmental stressors: The mental health impacts of living near industrial activity. *Journal of Health and Social Behavior*, *46*(3), 289–305. https://doi.org/10.1177/002214650504600306

Eimer, M. (2014). The neural basis of attentional control in visual search. *Trends in Cognitive Sciences*, *18*(10), 526–535. https://doi.org/10.1016/j.tics.2014.05.005

Eimontaite, I., Voinescu, A., Alford, C., Caleb-Solly, P., & Morgan, P. (2020). The Impact of Different Human-Machine Interface Feedback Modalities on Older Participants User Experience of CAVs in a Simulator Environment. *Advances in Intelligent Systems and Computing*, 120–132. https://doi.org/10.1007/978-3-030-20503-4_11

Endsley, M. R. (1996). Automation and situation awareness. *Automation and Human Performance: Theory and Applications*, *20*, 163–181. https://doi.org/10.1201/9781315137957

Endsley, M. R., & Garland, D. J. (2000). Pilot Situation Awareness Training in General Aviation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 357–360. https://doi.org/10.1177/154193120004401107

Esco, M. R., & Flatt, A. A. (2014). Ultra-short-term heart rate variability indexes at rest and post-exercise in athletes: Evaluating the agreement with accepted recommendations. *Journal of Sports Science and Medicine*, *13*(3), 535–541.

Esterman, M., Grosso, M., Liu, G., Mitko, A., Morris, R., & DeGutis, J. (2016). Anticipation of monetary reward can attenuate the vigilance decrement. *PLoS ONE*, *11*(7), 1–19. https://doi.org/10.1371/journal.pone.0159741

Esterman, M., Noonan, S. K., Rosenberg, M., & Degutis, J. (2013). In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cerebral Cortex*, *23*(11), 2712–2723. https://doi.org/10.1093/cercor/bhs261

Esterman, M., & Rothlein, D. (2019). Models of sustained attention. *Current Opinion in Psychology*, *29*, 174–180. https://doi.org/10.1016/j.copsyc.2019.03.005

Faber, K., & van Lierop, D. (2020). How will older adults use automated vehicles? Assessing the role of AVs in overcoming perceived mobility barriers. *Transportation Research Part A: Policy and Practice*, *133*, 353–363. https://doi.org/10.1016/j.tra.2020.01.022

Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, *77*, 167–181. https://doi.org/10.1016/j.tra.2015.04.003

Fairclough, S. H., & Houston, K. (2004). A metabolic measure of mental effort. *Biological Psychology*, *66*(2), 177–190. https://doi.org/10.1016/j.biopsycho.2003.10.001

Fan, J., Gu, X., Guise, K. G., Liu, X., Fossella, J., Wang, H., & Posner, M. I. (2009). Testing the behavioral interaction and integration of attentional networks. *Brain and Cognition*, *70*(2), 209–220. https://doi.org/10.1016/j.bandc.2009.02.002

Feldhutter, A., Kroll, D., & Bengler, K. (2018). Wake up and take over! The effect of fatigue on the take-over performance in conditionally automated driving. *21st International Conference on Intelligent Transportation Systems*, 2080–2085. https://doi.org/10.1109/ITSC.2018.8569545

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Thousand Oaks, CA: Sage Publications.

Finkelman, J. M. (1994). A large database study of the factors associated with work-induced fatigue. *Human Factors*, *36*(2), 232–243. https://doi.org/10.1177/001872089403600205

Flach, J. M. (1995). Situation awareness: Proceed with caution. *Human Factors*, *37*(1), 149–157. https://doi.org/10.1518/001872095779049480

Fortenbaugh, F. C., Degutis, J., & Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research. *Annals of the New York Academy of Sciences*, *1396*(1), 70–91. https://doi.org/10.1111/nyas.13318.Recent

Fortin, P. E., Sulmont, E., & Cooperstock, J. (2019). Detecting Perception of Smartphone Notifications using Skin Conductance Responses. *Conference on Human Factors in Computing Systems Proceedings*, 1–9. https://doi.org/10.1145/3290605.3300420

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(27), 9673–9678. https://doi.org/10.1073/pnas.0504136102

Foxe, J. J., Murphy, J. W., & De Sanctis, P. (2014). Throwing out the rules: Anticipatory alpha-band oscillatory attention mechanisms during task-set reconfigurations. *European Journal of Neuroscience*, *39*(11), 1960–1972. https://doi.org/10.1111/ejn.12577

Foxe, J. J., Simpson, G. V., & Ahlfors, S. P. (1998). Parieto-occipital ~10 Hz activity reflects anticipatory state of visual attention mechanisms. *NeuroReport*, *9*(17), 3929–3933. https://doi.org/10.1097/00001756-199812010-00030

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, *2*, 1–13. https://doi.org/10.3389/fpsyg.2011.00154

Friese, U., Daume, J., Göschl, F., König, P., Wang, P., & Engel, A. K. (2016). Oscillatory brain activity during multisensory attention reflects activation, disinhibition, and cognitive control. *Scientific Reports*, *6*, 32775. https://doi.org/10.1038/srep32775

Fuentes-García, J. P., Villafaina, S., Collado-Mateo, D., de la Vega, R., Olivares, P. R., & Clemente-Suárez, V. J. (2019). Differences between high vs. low performance chess players in heart rate variability during chess problems. *Frontiers in Psychology*, *10*, 1–9. https://doi.org/10.3389/fpsyg.2019.00409

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis and Prevention*, *37*(3), 461–472. https://doi.org/10.1016/j.aap.2004.11.003

Gable, P., & Harmon-Jones, E. (2010). The blues broaden, but the nasty narrows: Attentional consequences of negative affects low and high in motivational intensity. *Psychological Science*, *21*(2), 211–215. https://doi.org/10.1177/0956797609359622

Gagl, B., Hawelka, S., & Hutzler, F. (2011). Systematic influence of gaze position on pupil size measurement: Analysis and correction. *Behavior Research Methods*, *43*(4), 1171–1181. https://doi.org/10.3758/s13428-011-0109-5

Galante, F., Bracco, F., Chiorri, C., Pariota, L., Biggero, L., & Bifulco, G. N. (2018). Validity of mental workload measures in a driving simulation environment. *Journal of Advanced Transportation*, *2018*, 1–11. https://doi.org/10.1155/2018/5679151

Gambi, E., Agostinelli, A., Belli, A., Burattini, L., Cippitelli, E., Fioretti, S., Pierleoni, P., Ricciuti, M., Sbrollini, A., & Spinsante, S. (2017). Heart Rate Detection Using Microsoft Kinect: Validation and Comparison to Wearable Devices. *Sensors*, *17*(8), 1776. https://doi.org/10.3390/s17081776

Gao, M., Kortum, P., & Oswald, F. (2018). Psychometric evaluation of the USE (usefulness, satisfaction, and ease of use) questionnaire for reliability and validity. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1414–1418. https://doi.org/10.1177/1541931218621322

Gaume, A., Dreyfus, G., & Vialatte, F. B. (2019). A cognitive brain–computer interface monitoring sustained attentional variations during a continuous task. *Cognitive Neurodynamics*, *13*(3), 257–269. https://doi.org/10.1007/s11571-019-09521-4

Gevins, A., & Cutillo, B. (1993). Spatiotemporal dynamics of component processes in human working memory. *Electroencephalography and Clinical Neurophysiology*, *87*(3), 128–143. https://doi.org/10.1016/0013-4694(93)90119-G

Giakoumaki, S. G., Roussos, P., Pallis, E. G., & Bitsios, P. (2011). Sustained attention and working memory deficits follow a familial pattern in schizophrenia. *Archives of Clinical Neuropsychology*, *26*(7), 687–695. https://doi.org/10.1093/arclin/acr060

Giang, W. C. W., Shanti, I., Chen, H.-Y. W., Zhou, A., & Donmez, B. (2015). Smartwatches vs. smartphones: a preliminary report of driver behavior and perceived risk while responding to notifications. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 154–161. https://doi.org/10.1145/2799250.2799282

Giraudet, L., Imbert, J. P., Bérenger, M., Tremblay, S., & Causse, M. (2015). The neuroergonomic evaluation of human machine interface design in air traffic control using behavioral and EGG/ERP measures. *Behavioural Brain Research*, *294*, 246–253. https://doi.org/10.1016/j.bbr.2015.07.041

Glass, A. L., & Kang, M. (2019). Dividing attention in the classroom reduces exam performance. *Educational Psychology*, *39*(3), 395–408. https://doi.org/10.1080/01443410.2018.1489046

Gmehlin, D., Fuermaier, A. B. M., Walther, S., Tucha, L., Koerts, J., Lange, K. W., Tucha, O., Weisbrod, M., & Aschenbrenner, S. (2016). Attentional lapses of adults with attention deficit hyperactivity disorder in tasks of sustained attention. *Archives of Clinical Neuropsychology*, *31*(4), 343–357. https://doi.org/10.1093/arclin/acw016

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

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

Gold, C., Körber, M., Hohenberger, C., Lechner, D., & Bengler, K. (2015). Trust in Automation – Before and After the Experience of Take-over Scenarios in a Highly Automated Vehicle. *Procedia Manufacturing*, *3*, 3025–3032. https://doi.org/10.1016/j.promfg.2015.07.847

Gold, C., Körber, M., Lechner, D., & Bengler, K. (2016). Taking over Control from Highly Automated Vehicles in Complex Traffic Situations. *Human Factors*, *58*(4), 642–652. https://doi.org/10.1177/0018720816634226

Gonçalves, J. S., Moriguchi, C. S., Takekawa, K. S., Gil Coury, H. J. C., & Sato, T. D. O. (2017). The effects of forearm support and shoulder posture on upper trapezius and anterior deltoid activity. *Journal of Physical Therapy Science*, *29*(5), 793–798. https://doi.org/10.1589/jpts.29.793

Gordan, R., Gwathmey, J. K., & Xie, L.-H. (2015). Autonomic and endocrine control of cardiovascular function. *World Journal of Cardiology*, *7*(4), 204–214. https://doi.org/10.4330/wjc.v7.i4.204

Grefkes, C., Wang, L. E., Eickhoff, S. B., & Fink, G. R. (2010). Noradrenergic modulation of cortical networks engaged in visuomotor processing. *Cerebral Cortex*, *20*(4), 783–797. https://doi.org/10.1093/cercor/bhp144

Greicius, M. D., & Menon, V. (2004). Default-mode activity during a passive sensory task: Uncoupled from deactivation but impacting activation. *Journal of Cognitive Neuroscience*, *16*(9), 1484–1492. https://doi.org/10.1162/0898929042568532

Gruden, T., Stojmenova, K., Sodnik, J., & Jakus, G. (2019). Assessing drivers' physiological responses using consumer grade devices. *Applied Sciences*, *9*, 5353. https://doi.org/10.3390/app9245353

Grupe, D., & Nitschke, J. (2013). Uncertainty and Anticipation in Anxiety. *Nature Reviews Neuroscience*, *14*(7), 488–501. https://doi.org/10.1038/jid.2014.371

Gunnar, M., & Quevedo, K. (2007). The Neurobiology of Stress and Development. *Annual Review of Psychology*, *58*, 145–173. https://doi.org/10.1146/annurev.psych.58.110405.085605

Gunzelmann, G., Moore, L., Salvucci, D. D., & Gluck, K. A. (2011). Sleep loss and driver performance: Quantitative predictions with zero free parameters. *Cognitive Systems Research*, *12*(2), 154–163. https://doi.org/10.1016/j.cogsys.2010.07.009

Guo, Y., Wang, X., Xu, Q., Liu, F., Liu, Y., & Xia, Y. (2019). Change-point analysis of eye movement characteristics for female drivers in anxiety. *International Journal of Environmental Research and Public Health*, *16*(7), 1236. https://doi.org/10.3390/ijerph16071236

Haegens, S., Cousijn, H., Wallis, G., Harrison, P. J., & Nobre, A. C. (2014). Inter- and intra-individual variability in alpha peak frequency. *NeuroImage*, *92*, 46–55. https://doi.org/10.1016/j.neuroimage.2014.01.049

Haegens, S., Händel, B. F., & Jensen, O. (2011). Top-down controlled alpha band activity in somatosensory areas determines behavioral performance in a discrimination task. *Journal of Neuroscience*, *31*(14), 5197–5204. https://doi.org/10.1523/JNEUROSCI.5199-10.2011

Han, S. W., & Marois, R. (2014). Functional fractionation of the stimulus-driven attention network. *Journal of Neuroscience*, *34*(20), 6958–6969. https://doi.org/10.1523/JNEUROSCI.4975-13.2014

Harding, I. H., Yücel, M., Harrison, B. J., Pantelis, C., & Breakspear, M. (2015). Effective connectivity within the frontoparietal control network differentiates cognitive control and working memory. *NeuroImage*, *106*, 144–153. https://doi.org/10.1016/j.neuroimage.2014.11.039

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 904–908.

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

Head, J., & Helton, W. S. (2014). Sustained attention failures are primarily due to sustained cognitive load not task monotony. *Acta Psychologica*, *153*, 87–94. https://doi.org/10.1016/j.actpsy.2014.09.007

Hebb, D. O. (1955). Drives and the C. N. S. (conceptual nervous system). *Psychological Review*, *62*(4), 243–254. https://doi.org/10.1037/h0041823

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience*, *8*, 150. https://doi.org/10.3389/fnins.2014.00150

Helldin, T., Falkman, G., Riveiro, M., & Davidsson, S. (2013). Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 210–217. https://doi.org/10.1145/2516540.2516554

Henderson, J. M., Pollatsek, A., & Rayner, K. (1987). Effects of Foveal Priming and Extrafoveal

Preview on Object Identification. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(3), 449–463. https://doi.org/10.1037/0096-1523.13.3.449

Hergeth, S., Lorenz, L., Vilimek, R., & Krems, J. F. (2016). Keep Your Scanners Peeled: Gaze Behavior as a Measure of Automation Trust during Highly Automated Driving. *Human Factors*, *58*(3), 509–519. https://doi.org/10.1177/0018720815625744

Hermens, H. J., Freriks, B., Disselhorst-Klug, C., & Rau, G. (2000). Development of recommendations for SEMG sensors and sensor placement procedures. *Journal of Electromyography and Kinesiology*, *10*(5), 361–374. https://doi.org/10.1016/S1050-6411(00)00027-4

Herweg, N. A., Solomon, E. A., & Kahana, M. J. (2020). Theta Oscillations in Human Memory. *Trends in Cognitive Sciences*, *24*(3), 208–227. https://doi.org/10.1016/j.tics.2019.12.006

Hidalgo-Muñoz, A. R., Jallais, C., Evennou, M., Ndiaye, D., Moreau, F., Ranchet, M., Derollepot, R., & Fort, A. (2019). Hemodynamic responses to visual cues during attentive listening in autonomous versus manual simulated driving: A pilot study. *Brain and Cognition*, *135*, 103583. https://doi.org/10.1016/j.bandc.2019.103583

Hitchcock, E. M., Dember, W. N., Warm, J. S., Moroney, B. W., & See, J. E. (1999). Effects of cueing and knowledge of results on workload and boredom in sustained attention. *Human Factors*, *41*(3), 363–372. https://doi.org/10.1518/001872099779610987

Ho, C., Gray, R., & Spence, C. (2014). To what extent do the findings of laboratory-based spatial attention research apply to the real-world setting of driving? *IEEE Transactions on Human-Machine Systems*, *44*(4), 524–530. https://doi.org/10.1109/THMS.2014.2316502

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods*, *39*, 101–117. https://doi.org/10.3758/BF03192848

Hommel, B., Johnston, J., Logan, G., & Pashler, H. (1998). Automatic stimulus-response translation in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(5), 1368–1384.

Horstmann, G., & Herwig, A. (2016). Novelty biases attention and gaze in a surprise trial. *Attention, Perception, and Psychophysics*, *78*(1), 69–77. https://doi.org/10.3758/s13414-015-0995-1

Huang, R. S., Jung, T. P., Delorme, A., & Makeig, S. (2007). Tonic and phasic electroencephalographic dynamics during continuous compensatory tracking. *NeuroImage*, *39*, 1896–1909. https://doi.org/10.1016/j.neuroimage.2007.10.036

Huang, R. S., Jung, T. P., & Makeig, S. (2009). Tonic changes in EEG power spectra during simulated driving. *International Conference on Foundations of Augmented Cognition*, 394–403. https://doi.org/10.1007/978-3-642-02812-0_47

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

Imbert, J. P., Hodgetts, H. M., Parise, R., Vachon, F., Dehais, F., & Tremblay, S. (2014). Attentional costs and failures in air traffic control notifications. *Ergonomics*, *57*(12), 1817–1832. https://doi.org/10.1080/00140139.2014.952680

Ivry, R., & Robertson, L. C. (1998). *The Two Sides of Perception*. Cambridge, MA: MIT Press. https://doi.org/10.1016/s1364-6613(98)01165-6

Jacobé, C., Naurois, D., Bourdin, C., Stratulat, A., Diaz, E., & Vercher, J. (2019). Detection and prediction of driver drowsiness using artificial neural network models. *Accident Analysis and Prevention*, *126*, 95–104. https://doi.org/10.1016/j.aap.2017.11.038

Jamadar, S. D., Thienel, R., & Karayanidis, F. (2015). Task Switching Processes. *Brain Mapping: An Encyclopedic Reference*, *1*, 327–335. https://doi.org/10.1016/B978-0-12-397025-1.00250-5

Jarosch, O., Bellem, H., & Bengler, K. (2019). Effects of Task-Induced Fatigue in Prolonged Conditional Automated Driving. *Human Factors*, *61*(7), 1186–1199. https://doi.org/10.1177/0018720818816226

Jayasinghe, S. U., Torres, S. J., Nowson, C. A., Tilbrook, A. J., & Turner, A. I. (2014). Physiological responses to psychological stress: importance of adiposity in men aged 50–70 years. *Endocrine Connections*, *3*(3), 110–119. https://doi.org/10.1530/ec-14-0042

Jensen, O., Gelfand, J., Kounios, J., & Lisman, J. E. (2002). Oscillations in the alpha band (9-12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex*, *12*(8), 877–882. https://doi.org/10.1093/cercor/12.8.877

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience*, *4*(186), 1–8. https://doi.org/10.3389/fnhum.2010.00186

Joshi, S., Li, Y., Kalwani, R., & Gold, J. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 221–234. https://doi.org/10.1016/j.physbeh.2017.03.040

Kam, J. W. Y., Solbakk, A. K., Funderud, I., Endestad, T., Meling, T. R., & Knight, R. T. (2018). Orbitofrontal damage reduces auditory sensory response in humans. *Cortex*, *101*, 309–312. https://doi.org/10.1016/j.cortex.2017.12.023

Kang, H. G., Mahoney, D. F., Hoenig, H., Hirth, V. A., Bonato, P., Hajjar, I., & Lipsitz, L. A. (2010). In situ monitoring of health in older adults: Technologies and issues. *Journal of the American Geriatrics Society*, *58*(8), 1579–1586. https://doi.org/10.1111/j.1532-5415.2010.02959.x

Kantowitz, B. H., Becker, C. A., & Barlow, T. S. (1993). Assessing driver acceptance of IVHS components. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1062–1066. https://doi.org/10.1177/154193129303701605

Kappenman, E. S., & Luck, S. J. (2012). *The Oxford Handbook of Event-Related Potential Components*. New York, NY: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.001.0001

Kappenman, E. S., & Luck, S. J. (2016). Best Practices for Event-Related Potential Research in Clinical Populations. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *1*(2), 110–115. https://doi.org/10.1016/j.bpsc.2015.11.007

Karthaus, M., Wascher, E., & Getzmann, S. (2018). Effects of Visual and Acoustic Distraction on Driving Behavior and EEG in Young and Older Car Drivers: A Driving Simulation Study. *Frontiers in Aging Neuroscience*, *10*, 420. https://doi.org/10.3389/fnagi.2018.00420

Kassner, A., Muhrer, E., Baumann, M., Vollrath, M., Minin, L., Megard, C., & Heers, R. (2011). *Integrated human modelling and simulation to support human error risk analysis of partially autonomous driver assistance systems. Knowledge-base of driving without and with PADAS (Attachment to Deliverable D1.3)*. Available Online at: Https://Www.Semanticscholar.Org/Paper/Integrated-Human-Modelling-and-Simulation-to-Human-Aleixandre-Pena/6ca21a0c9c6519bf6ad5811164572d2563561753 (Accessed Sep 30 2020).

Kassner, L., Hirmer, P., Wieland, M., Steimle, F., Königsberger, J., & Mitschang, B. (2017). The Social Factory: Connecting People, Machines and Data in Manufacturing for Context-Aware Exception Escalation. *Proceedings of the 50th Hawaii International Conference on System Sciences*, 1673–1682. https://doi.org/10.24251/hicss.2017.202

Käthner, I., Wriessnegger, S. C., Müller-Putz, G. R., Kübler, A., & Halder, S. (2014). Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain-computer interface. *Biological Psychology*, *102*(1), 118–129. https://doi.org/10.1016/j.biopsycho.2014.07.014

Kelley, J. F. (1985). CAL - a natural language program developed with the OZ paradigm: Implications for supercomputing systems. *First International Conference on Supercomputing Systems*, 238–248.

Kharrazi, S., Augusto, B., & Fröjd, N. (2020). Vehicle dynamics testing in motion based driving simulators. *Vehicle System Dynamics*, *58*(1), 92–107.

https://doi.org/10.1080/00423114.2019.1566555

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching-a review. *Psychological Bulletin*, *136*(5), 849–874. https://doi.org/10.1037/a0019842

Kim, C., Johnson, N. F., & Gold, B. T. (2012). Common and distinct neural mechanisms of attentional switching and response conflict. *Brain Research*, *1469*, 92–102. https://doi.org/10.1016/j.brainres.2012.06.013

Klauck, M., Sugano, Y., & Bulling, A. (2017). Noticeable or distractive? A design space for gaze-contingent user interface notifications. *Conference on Human Factors in Computing Systems - Proceedings*, 1779–1786. https://doi.org/10.1145/3027063.3053085

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

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. https://doi.org/10.1016/j.tics.2012.10.007

Klimesch, W., Sauseng, P., & Gerloff, C. (2003). Enhancing cognitive performance with repetitive transcranial magnetic stimulation at human individual alpha frequency. *European Journal of Neuroscience*, *17*(5), 1129–1133. https://doi.org/10.1046/j.1460-9568.2003.02517.x

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Ko, S. M., & Ji, Y. G. (2018). How we can measure the non-driving-task engagement in automated driving: Comparing flow experience and workload. *Applied Ergonomics*, *67*, 237–245. https://doi.org/10.1016/j.apergo.2017.10.009

Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking-an integrative review of dual-task and task-switching research. *Psychological Bulletin*, *144*(6), 557–583. https://doi.org/10.1037/bul0000144

Körber, M., Baseler, E., & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics*, *66*, 18–31. https://doi.org/10.1016/j.apergo.2017.07.006

Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance decrement and passive fatigue caused by monotony in automated driving. *Procedia Manufacturing*, *3*, 2403–2409. https://doi.org/10.1016/j.promfg.2015.07.499

Körner, U., Müller-Thur, K., Lunau, T., Dragano, N., Angerer, P., & Buchner, A. (2019). Perceived stress in human–machine interaction in modern manufacturing environments—Results of a qualitative interview study. *Stress and Health*, *35*(2), 187–199. https://doi.org/10.1002/smi.2853

Kowler, E. (2011). Eye movements: The past 25 years Eileen. *Vision Research*, *51*(13), 1457–1483. https://doi.org/10.1016/j.visres.2010.12.014.Eye

Kramer, A. F., & Madden, D. (2008). Attention. In *F. I. M. Craik & T. A. Salthouse (Eds.), The Handbook of Aging and Cognition* (pp. 189–249). Hillsdale, NJ: Lawrence Erlbaum.

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(6), 661–679. https://doi.org/10.1017/S0140525X12003196

Kwon, O., Jeong, J., Kim, H. Bin, Kwon, I. H., Park, S. Y., Kim, J. E., & Choi, Y. (2018). Electrocardiogram sampling frequency range acceptable for heart rate variability analysis. *Healthcare Informatics Research*, *24*(3), 198–206. https://doi.org/10.4258/hir.2018.24.3.198

Laborde, S., Mosley, E., & Thayer, J. F. (2017). Heart rate variability and cardiac vagal tone in psychophysiological research - Recommendations for experiment planning, data analysis, and data reporting. *Frontiers in Psychology*, *8*, 1–18. https://doi.org/10.3389/fpsyg.2017.00213

Lakens, D. (2017). Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. https://doi.org/10.1177/1948550617697177

Lang, A., Gapenne, O., & Rovira, K. (2011). Questioning implicit motor learning as instantiated by the pursuit-tracking task. *Quarterly Journal of Experimental Psychology*, *64*(10), 2003–2011. https://doi.org/10.1080/17470218.2011.573566

Langner, R., & Eickhoff, S. B. (2013). Sustaining Attention to Simple Tasks: A Meta-Analytic Review of the Neural Mechanisms of Vigilant Attention. *Psychol Bull*, *139*(4), 870–900. https://doi.org/10.1038/jid.2014.371

Larsson, A. F. L., Kircher, K., & Hultgren, J. A. (2014). Learning from experience: Familiarity with ACC and responding to a cut-in situation in automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*(B), 229–237. https://doi.org/10.1016/j.trf.2014.05.008

Laumann, K., Gärling, T., & Stormak, K. M. (2003). Selective attention and heart rate responses to natural and urban environments. *Journal of Environmental Psychology*, *23*(2), 125–134. https://doi.org/10.1016/S0272-4944(02)00110-X

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in*

*Psychological Science*, *19*(3), 143–148. https://doi.org/10.1177/0963721410370295

Lavie, N., & Dalton, P. (2014). Load Theory of Attention and Cognitive Control. In *A. C. Nobre & S. Kastner (Eds.), The Oxford Handbook of Attention* (pp. 56–75). New York, NY: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199675111.013.003

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*, 183–197. https://doi.org/10.3758/BF03213897

Lee, H. (2002). The Validity of Driving Simulator to Measure On-road Driving Performance of Older Drivers. *Transport Engineering in Australia*, *8*(2), 1–14.

Lehrer, P. (2013). How Does Heart Rate Variability Biofeedback Work? Resonance, the Baroreflex, and Other Mechanisms. *Biofeedback*, *5*, 756. https://doi.org/10.5298/1081-5937-41.1.02

Lenartowicz, A., Simpson, G. V., & Cohen, M. S. (2013). Perspective: Causes and functional significance of temporal variations in attention control. *Frontiers in Human Neuroscience*, *7*, 381. https://doi.org/10.3389/fnhum.2013.00381

Li, K. Z. H., Krampe, R. ., & Bondar, A. (2005). An ecological approach to studying aging and dual-task performance. In *R. W. Engle, G. Sedek, U. von Hecker, & D. N. McIntosh (Eds.), Cognitive limitations in aging and psychopathology* (pp. 190–218). New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511720413.009

Li, L., Wang, M., Zhao, Q., & Fogelson, N. (2012). Neural Mechanisms Underlying the Cost of Task Switching: An ERP Study. *PLoS ONE*, *7*(7), e42233. https://doi.org/10.1371/journal.pone.0042233

Li, S., Blythe, P., Guo, W., & Namdeo, A. (2018). Investigation of older driver's takeover performance in highly automated vehicles in adverse weather conditions. *IET Intelligent Transport Systems*, *12*(9), 1157–1165. https://doi.org/10.1049/iet-its.2018.0104

Liang, N., & Pitts, B. J. (2019). The Effect of Secondary Cognitive Task Difficulty on Headway Maintenance and Perceived Workload While Using Lane Keeping Systems. *Proceedings of the Human Factors and Ergonomics Society 2019 Annual Meeting*, *63*, 2011–2015. https://doi.org/10.1177/1071181319631161

Liang, Y., Reyes, M. L., & Lee, J. D. (2007). Real-Time Detection of Driver Cognitive Distraction Using Support Vector Machines. *IEEE Transactions on Intelligent Transportation Systems*, *8*(2), 340–349. https://doi.org/10.1109/TITS.2007.895298

Lin, L., Cockerham, D., Chang, Z., & Natividad, G. (2016). Task Speed and Accuracy Decrease When Multitasking. *Technology, Knowledge and Learning*, *21*(3), 307–323. https://doi.org/10.1007/s10758-015-9266-4

Lin, P., Yang, Y., Gao, J., De Pisapia, N., Ge, S., Wang, X., Zuo, C. S., Levitt, J., & Niu, C. (2017).

Dynamic Default Mode Network across Different Brain States. *Scientific Reports*, *7*, 46088. https://doi.org/10.1038/srep46088

Linden, D. E. J. (2005). The P300: Where in the brain is it produced and what does it tell us? *Neuroscientist*, *11*(6), 563–576. https://doi.org/10.1177/1073858405280524

Linnell, K. J., & Caparos, S. (2011). Perceptual and Cognitive Load Interact to Control the Spatial Focus of Attention. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1643. https://doi.org/10.1037/a0024669

Liston, C., Matalon, S., Hare, T. A., Davidson, M. C., & Casey, B. J. (2006). Anterior Cingulate and Posterior Parietal Cortices Are Sensitive to Dissociable Forms of Conflict in a Task-Switching Paradigm. *Neuron*, *50*(4), 643–653. https://doi.org/10.1016/j.neuron.2006.04.015

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*(57), 1–27. https://doi.org/10.3389/fnhum.2019.00057

Lorah, J. (2018). Effect size measures for multilevel models: definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education*, *6*(8), 1–11. https://doi.org/10.1186/s40536-018-0061-2

Louw, T., & Merat, N. (2017). Are you in the loop? Using gaze dispersion to understand driver visual attention during vehicle automation. *Transportation Research Part C: Emerging Technologies*, *76*, 35–50. https://doi.org/10.1016/j.trc.2017.01.001

Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability Interface*, *8*(2), 3–6. https://doi.org/10.1177/1078087402250360

Luque-Casado, A., Perales, J. C., Cárdenas, D., & Sanabria, D. (2016). Heart rate variability and cognitive processing: The autonomic response to task demands. *Biological Psychology*, *113*, 83–90. https://doi.org/10.1016/j.biopsycho.2015.11.013

Luque-Casado, A., Zabala, M., Morales, E., Mateo-March, M., & Sanabria, D. (2013). Cognitive Performance and Heart Rate Variability: The Influence of Fitness Level. *PLoS ONE*, *8*(2), e56935. https://doi.org/10.1371/journal.pone.0056935

Ma, J., Gu, J., Jia, H., Yao, Z., & Chang, R. (2018). The Relationship Between Drivers ’ Cognitive Fatigue and Speed Variability During Monotonous Daytime Driving. *Frontiers in Psychology*, *9*, 459. https://doi.org/10.3389/fpsyg.2018.00459

Macdonald, J. S. P., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, and Psychophysics*, *73*(6), 1780–1789. https://doi.org/10.3758/s13414-011-0144-4

Mackenzie, A. K., & Harris, J. M. (2015). Eye movements and hazard perception in active and passive driving. *Visual Cognition*, *23*(6), 736–757. https://doi.org/10.1080/13506285.2015.1079583

Mackworth, N. H. (1948). The Breakdown of Vigilance during Prolonged Visual Search. *Quarterly Journal of Experimental Psychology*, *1*, 6–21. https://doi.org/10.1080/17470214808416738

MacLean, K. A., Ferrer, E., Aichele, S. R., Bridwell, D. A., Zanesco, A. P., Jacobs, T. L., King, B. G., Rosenberg, E. L., Sahdra, B. K., Shaver, P. R., Wallace, B. A., Mangun, G. R., & Saron, C. D. (2010). Intensive meditation training improves perceptual discrimination and sustained attention. *Psychological Science*, *21*(6), 829–839. https://doi.org/10.1177/0956797610371339

Madigan, R., Louw, T., & Merat, N. (2018). The effect of varying levels of vehicle automation on drivers' lane changing behaviour. *PLoS ONE*, *13*(2), e0192190. https://doi.org/10.1371/journal.pone.0192190

Magosso, E., De Crescenzio, F., Ricci, G., Piastra, S., & Ursino, M. (2019). EEG alpha power is modulated by attentional changes during cognitive tasks and virtual reality immersion. *Computational Intelligence and Neuroscience*, 7051079. https://doi.org/10.1155/2019/7051079

Mandrick, K., Peysakhovich, V., Rémy, F., Lepron, E., & Causse, M. (2016). Neural and psychophysiological correlates of human performance under stress and high mental workload. *Biological Psychology*, *121*, 62–73. https://doi.org/10.1016/j.biopsycho.2016.10.002

Mann, W. C., Belchior, P., Tomita, M. R., & Kemp, B. J. (2007). Older adults' perception and use of PDAs, home automation system, and home health monitoring system. *Topics in Geriatric Rehabilitation*, *23*(1), 35–46. https://doi.org/10.1097/00013614-200701000-00006

Marker, R. J., Campeau, S., & Maluf, K. S. (2017). Psychosocial stress alters the strength of reticulospinal input to the human upper trapezius. *Journal of Neurophysiology*, *117*(1), 457–466. https://doi.org/10.1152/jn.00448.2016

Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, *9*(6), 296–305. https://doi.org/10.1016/j.tics.2005.04.010

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, *5*, 229–240. https://doi.org/10.1038/nrn1348

Mayr, U., Kuhns, D., & Hubbard, J. (2014). Long-term memory and the control of attentional control. *Cognitive Psychology*, *72*, 1–26. https://doi.org/10.1016/j.cogpsych.2014.02.001

McBride, S. E., Rogers, W. A., & Fisk, A. D. (2011). Understanding the effect of workload on automation use for younger and older adults. *Human Factors*, *53*(6), 672–686. https://doi.org/10.1177/0018720811421909

McBride, S. E., Rogers, W. A., & Fisk, A. D. (2010). Do younger and older adults differentially depend on an automated system? *Proceedings of the Human Factors and Ergonomics Society Annual*

*Meeting*, 175–179. https://doi.org/10.1518/107118110X12829369200233

McCall, R., McGee, F., Mirnig, A., Meschtscherjakov, A., Louveton, N., Engel, T., & Tscheligi, M. (2019). A taxonomy of autonomous vehicle handover situations. *Transportation Research Part A: Policy and Practice*, *124*, 507–522. https://doi.org/10.1016/j.tra.2018.05.005

McKendrick, R., & Harwood, A. (2019). Cognitive Workload and Workload Transitions Elicit Curvilinear Hemodynamics During Spatial Working Memory. *Frontiers in Human Neuroscience*, *13*, 1–16. https://doi.org/10.3389/fnhum.2019.00405

McKewen, M., Cooper, P. S., Wong, A. S. W., Michie, P. T., Sauseng, P., & Karayanidis, F. (2020). Task-switching costs have distinct phase-locked and nonphase-locked EEG power effects. *Psychophysiology*, *57*(5), e13533. https://doi.org/10.1111/psyp.13533

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

Meinlschmidt, G., Stalujanis, E., & Tegethoff, M. (2018). The psychobiology of using automated driving systems: A systematic review and integrative model. *Psychoneuroendocrinology*, *105*, 1–13. https://doi.org/10.1016/j.psyneuen.2018.09.029

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning Memory and Cognition*, *22*(6), 1423. https://doi.org/10.1037/0278-7393.22.6.1423

Melnicuk, V., Birrell, S., Crundall, E., & Jennings, P. (2016). Towards hybrid driver state monitoring: Review, future perspectives and the role of consumer electronics. *IEEE Intelligent Vehicles Symposium,* 1392–1397. https://doi.org/10.1109/IVS.2016.7535572

Melnicuk, V., Birrell, S., Thompson, S., Mouzakitis, A., & Jennings, P. (2019). How Acceptable Is It to Monitor Driver State? Using the UTAUT Model to Analyse Drivers' Perceptions Towards the System. *Advances in Intelligent Systems and Computing*, 579–890. https://doi.org/10.1007/978-3-319-93885-1_53

Menghini, L., Gianfranchi, E., Cellini, N., Patron, E., Tagliabue, M., & Sarlo, M. (2019). Stressing the accuracy: Wrist-worn wearable sensor validation over different conditions. *Psychophysiology*, *56*(11), e13441. https://doi.org/10.1111/psyp.13441

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure & Function*, *214*(5–6), 655–667. https://doi.org/10.1007/s00429-010-0262-0

Merat, N., Jamson, A. H., Lai, F. C. H., Daly, M., & Carsten, O. M. J. (2014). Transition to manual:

Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*(B), 274–282. https://doi.org/10.1016/j.trf.2014.09.005

Mihailidis, A., Cockburn, A., Longley, C., & Boger, J. (2008). The acceptability of home monitoring technology among community-dwelling older adults and baby boomers. *Assistive Technology*, *20*(1), 1–12. https://doi.org/10.1080/10400435.2008.10131927

Missonnier, P., Deiber, M. P., Gold, G., Millet, P., Gex-Fabry Pun, M., Fazio-Costa, L., Giannakopoulos, P., & Ibáñez, V. (2006). Frontal theta event-related synchronization: Comparison of directed attention and working memory load effects. *Journal of Neural Transmission*, *113*(10), 1477–1486. https://doi.org/10.1007/s00702-005-0443-9

Mole, C. D., Lappi, O., Science, C., Mars, F., & Nantes, C. (2019). Getting Back Into the Loop : The Perceptual-Motor Determinants of Successful Transitions out of Automated Driving. *Human Factors*, *61*(7), 1037–1065. https://doi.org/10.1177/0018720819829594

Moray, N. (1959). Attention in Dichotic Listening: Affective Cues and the Influence of Instructions. *Quarterly Journal of Experimental Psychology*, *11*(1), 56–60. https://doi.org/10.1080/17470215908416289

Morgan, P. L., Voinescu, A., Williams, C., Caleb-Solly, P., Alford, C., Shergold, I., Parkhurst, G., & Pipe, A. (2018). An emerging framework to inform effective design of human-machine interfaces for older adults using connected autonomous vehicles. *Advances in Intelligent Systems and Computing*, 325–334. https://doi.org/10.1007/978-3-319-60441-1_33

Morris, D. M., Erno, J. M., & Pilcher, J. J. (2017). Electrodermal Response and Automation Trust during Simulated Self-Driving Car Use. *Proceedings of the Human Factors and Ergonomics Society 2017 Annual Meeting*, 1759–1762. https://doi.org/10.1177/1541931213601921

Mullen, N., Charlton, J., Devlin, A., & Bédard, M. (2011). Simulator validity: Behaviors observed on the simulator and on the road. In *D. Fisher, M. Rizzo, J. Caird, J. Lee (Eds.), Handbook of Driving Simulation for Engineering, Medicine, and Psychology* (pp. 1–18). Raton, FL: CRC Press.

Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory—Where are we now, and where should we go next? *Psychonomic Bulletin and Review*, *23*, 1316–1340. https://doi.org/10.3758/s13423-015-0982-5

Murta, T., Leite, M., Carmichael, D. W., Figueiredo, P., & Lemieux, L. (2015). Electrophysiological correlates of the BOLD signal for EEG-informed fMRI. *Human Brain Mapping*, *36*(1), 391–414. https://doi.org/10.1002/hbm.22623

Mutha, P. K., Haaland, K. Y., & Sainburg, R. L. (2013). Rethinking Motor Lateralization: Specialized but Complementary Mechanisms for Motor Control of Each Arm. *PLoS ONE*, *8*(3), e58582.

https://doi.org/10.1371/journal.pone.0058582

Muthukumaraswamy, S. D. (2013). High-frequency brain activity and muscle artifacts in MEG/EEG: A review and recommendations. *Frontiers in Human Neuroscience*, *7*, 1–11. https://doi.org/10.3389/fnhum.2013.00138

Nachreiner, F., & Hänecke, K. (1992). Vigilance. In *A. P. Smith & D. M. Jones (Eds.), Handbook of human performance, Vol. 1. The physical environment; Vol. 2. Health and performance; Vol. 3. State and trait* (pp. 261–288). Cambridge, MA: Academic Press. https://doi.org/10.1016/b978-0-12-650353-1.50016-2

Nascimento, A. M., Queiroz, A. C. M., Vismari, L. F., Bailenson, J. N., Cugnasca, P. S., Camargo, J. B., & De Almeida, J. R. (2019). The role of virtual reality in autonomous vehicles' safety. *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality*, 50–57. https://doi.org/10.1109/AIVR46125.2019.00017

Naujoks, F., Hö, S., Purucker, C., & Zeeb, K. (2018). From partial and high automation to manual driving: Relationship between non-driving related tasks, drowsiness and take-over performance. *Accident Analysis & Prevention*, *121*, 28–42. https://doi.org/10.1016/j.aap.2018.08.018

Nelson, J. T., McKinley, R. A., Golob, E. J., Warm, J. S., & Parasuraman, R. (2014). Enhancing vigilance in operators with prefrontal cortex transcranial direct current stimulation (tDCS). *NeuroImage*, *85*(3), 909–917. https://doi.org/10.1016/j.neuroimage.2012.11.061

Neubauer, C., Matthews, G., & Saxby, D. (2012). The effects of cell phone use and automation on driver performance and subjective state in simulated driving. *Proceedings of the Human Factors and Ergonomics Society*, *56*(1), 1987–1991. https://doi.org/10.1177/1071181312561415

Niedermeyer, E., & Lopes da Silva, F. H. (2005). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Philadelphia, PA: Lippincott Williams & Wilkins. https://doi.org/10.1212/01.wnl.0000243257.85592.9a

Nobre, A. C., & Mesulam, M. M. (2014). Large-scale networks for attentional biases. In *A. C. Nobre & S. Kastner (Eds.), The Oxford Handbook of Attention*. New York, NY: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199675111.013.035

Oken, B. S., Chamine, I., & Wakeland, W. (2015). A Systems Approach to Stress, Stressors and Resilience in Humans. *Behavioural Brain Research*, *282*, 144–154. https://doi.org/10.1016/j.bbr.2014.12.047

Oken, B. S., Salinsky, M. C., & Elsas, S. M. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology*, *117*(9), 1885–1901.

https://doi.org/10.1016/j.clinph.2006.01.017

Oliveira, F. T. P., Diedrichsen, J., Verstynen, T., Duque, J., & Ivry, R. B. (2010). Transcranial magnetic stimulation of posterior parietal cortex affects decisions of hand choice. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(41), 17751–17756. https://doi.org/10.1073/pnas.1006223107

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*(4), 713–719. https://doi.org/10.1016/S1388-2457(00)00527-7

Otero-Millan, J., Troncoso, X. G., Macknik, S. L., Serrano-Pedraza, I., & Martinez-Conde, S. (2008). Saccades and microsaccades during visual fixation, exploration, and search: Foundations for a common saccadic generator. *Journal of Vision*, *8*(21), 1–18. https://doi.org/10.1167/8.14.21

Ozimek, A., Lewandowska, P., Krejtz, K., & Duchowski, A. T. (2019). Attention towards privacy notifications on web pages. *Eye Tracking Research and Applications Symposium (ETRA)*, 1–5. https://doi.org/10.1145/3317960.3321618

Pacheco-Unguetti, A. P., Acosta, A., Callejas, A., & Lupiáñez, J. (2010). Attention and anxiety: Different attentional functioning under state and trait anxiety. *Psychological Science*, *21*(2), 298–304. https://doi.org/10.1177/0956797609359624

Pallant, J. (2007). *“Survival manual.” A step by step guide to data analysis using SPSS* (3rd ed.). New York, NY: McGraw-Hill.

Paquet, L., & Craig, G. L. (1997). Evidence for selective target processing with a low perceptual lead flankers task. *Memory and Cognition*, *25*, 182–189. https://doi.org/10.3758/BF03201111

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, *52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Parette, P., & Scherer, M. (2004). Assistive technology use and stigma. *Education and Training in Developmental Disabilities*, *39*(3), 217–226.

Park, J., & Park, S. H. (2007). The effects of various visual enhancements during continuous pursuit tracking tasks. *International Conference on Human-Computer Interaction*, 125–132.

Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 358–377. https://doi.org/10.1037/0096-1523.10.3.358

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*(2), 220–244. https://doi.org/10.1037/0033-2909.116.2.220

Paxion, J., Galy, E., & Berthelon, C. (2014). Mental workload and driving. *Frontiers in Psychology*, *5*, 1344. https://doi.org/10.3389/fpsyg.2014.01344

Payne, A. F. H., Schell, A. M., & Dawson, M. E. (2016). Lapses in skin conductance responding across anatomical sites: Comparison of fingers, feet, forehead, and wrist. *Psychophysiology*, *53*(7), 1084–1092. https://doi.org/10.1111/psyp.12643

Pelagatti, C., Binda, P., & Vannucci, M. (2018). Tracking the Dynamics of Mind Wandering: Insights from Pupillometry. *Journal of Cognition*, *1*(1), 38. https://doi.org/10.5334/joc.41

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*(2), 184–187. https://doi.org/10.1016/0013-4694(89)90180-6

Peters, M., Reimers, S., & Manning, J. T. (2006). Hand preference for writing and associations with selected demographic and behavioral variables in 255,100 subjects: The BBC internet study. *Brain and Cognition*, *62*, 177–189

Petersen, L., Robert, L., Yang, X. J., & Tilbury, D. (2019). Situational Awareness, Driver's Trust in Automated Driving Systems and Secondary Task Performance. *SAE International Journal of Connected and Automated Vehicles*, *2*(2), 1–13. https://doi.org/10.4271/12-02-02-0009

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

Pfurtscheller, G. (2006). The cortical activation model (CAM). *Progress in Brain Research*, *159*, 19–27. https://doi.org/10.1016/S0079-6123(06)59002-8

Pfurtscheller, G., Neuper, C., Andrew, C., & Edlinger, G. (1997). Foot and hand area mu rhythms. *International Journal of Psychophysiology*, *26*(1–3), 121–135. https://doi.org/10.1016/S0167-8760(97)00760-5

Phipps-Nelson, J. O., Redman, J. R., & Rajaratnam, S. M. W. (2011). Temporal profile of prolonged, night-time driving performance: Breaks from driving temporarily reduce time-on-task fatigue but not sleepiness. *Journal of Sleep Research*, *20*(3), 404–415. https://doi.org/10.1111/j.1365-2869.2010.00900.x

Polich, J. (1987). Task difficulty, probability, and inter-stimulus interval as determinants of P300 from auditory stimuli. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *68*(4), 311–320. https://doi.org/10.1016/0168-5597(87)90052-9

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Poljac, E., Koch, I., & Bekkering, H. (2009). Dissociating restart cost and mixing cost in task switching. *Psychological Research*, *73*(3), 407–416. https://doi.org/10.1007/s00426-008-0151-9

Pomplun, M., Garaas, T. W., & Carrasco, M. (2013). The effects of task difficulty on visual search strategy in virtual 3D displays. *Journal of Vision*, *13*(3), 1–22. https://doi.org/10.1167/13.3.24

Posada-Quintero, H. F., Florian, J. P., Orjuela-Cañón, A. D., Aljama-Corrales, T., Charleston-Villalobos, S., & Chon, K. H. (2016). Power Spectral Density Analysis of Electrodermal Activity for Sympathetic Function Assessment. *Annals of Biomedical Engineering*, *10*, 3124–3135. https://doi.org/10.1007/s10439-016-1606-6

Posner, M. I. (1975). Psychobiology of Attention. In *M. S. Gazzaniga & C. Blakemore (Eds.), Handbook of Psychobiology* (pp. 441–480). Cambridge, MA: Academic Press. https://doi.org/10.1016/b978-0-12-278656-3.50019-3

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Posner, M. I., & Petersen, S. E. (1990). The Attention System of the Human Brain. *Annual Review of Neuroscience*, *13*, 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Prinzel, L. J., Freeman, F. G., & Prinzel, H. D. (2005). Individual differences in complacency and monitoring for automation failures. *Individual Differences Research*, *3*(1), 27–49.

Proskovec, A. L., Heinrichs-Graham, E., Wiesman, A. I., McDermott, T. J., & Wilson, T. W. (2018). Oscillatory dynamics in the dorsal and ventral attention networks during the reorienting of attention. *Human Brain Mapping*, *39*(5), 2177–2190. https://doi.org/10.1002/hbm.23997

Proskovec, A. L., Wiesman, A. I., & Wilson, T. W. (2019). The strength of alpha and gamma oscillations predicts behavioral switch costs. *NeuroImage*, *188*, 274–281. https://doi.org/10.1016/j.neuroimage.2018.12.016

Puma, S., Matton, N., Paubel, P.-V., Raufaste, É., & El-Yagoubi, R. (2018). Using theta and alpha band power to assess cognitive workload in multitasking environments. *International Journal of Psychophysiology*, *123*, 111–120. https://doi.org/10.1016/J.IJPSYCHO.2017.10.004

Qin, W., Xiang, X., & Taylor, H. (2019). Driving Cessation and Social Isolation in Older Adults. *Journal of Aging and Health*, *29*, 1119–1143. https://doi.org/10.1177/0898264319870400

Quer, G., Gouda, P., Galarnyk, M., Topol, E. J., & Steinhubl, S. R. (2020). Inter- And intraindividual variability in daily resting heart rate and its associations with age, sex, sleep, BMI, and time of year: Retrospective, longitudinal cohort study of 92,457 adults. *PLoS ONE*, *15*(2), 1–12. https://doi.org/10.1371/journal.pone.0227709

Quintana, D. S., Guastella, A. J., Mcgregor, I. S., Hickie, I. B., & Kemp, A. H. (2013). Moderate alcohol intake is related to increased heart rate variability in young adults: Implications for health and well-being. *Psychophysiology*, *50*(12), 1202–1208. https://doi.org/10.1111/psyp.12134

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. *Proceedings of the Human Factors and Ergonomics Society*, *58*(1), 2063–2067. https://doi.org/10.1177/1541931214581434

Radlmayr, J., Ratter, M., Feldhütter, A., & Körber, M. (2019). Take-overs in Level 3 automated driving – proposal of the take-over performance score (TOPS). *Proceedings of the 20th Congress of the International Ergonomics Association. IEA 2018. Advances in Intelligent Systems and Computing*, *4*, 537–542. https://doi.org/10.1007/978-3-319-96074-6

Ragot, M., Martin, N., Em, S., Pallamin, N., & Diverrez, J.-M. (2017). Emotion recognition using physiological signals: Laboratory vs. wearable sensors. *Applied Human Factors and Ergonomics*, *48*, 813–822. https://doi.org/10.1007/978-3-319-60639-2_2

Raichle, M. E. (2015). The Brain's Default Mode Network. *Annual Review of Neuroscience*, *38*, 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030

Ralph, B. C. W., Onderwater, K., Thomson, D. R., & Smilek, D. (2017). Disrupting monotony while increasing demand: benefits of rest and intervening tasks on vigilance. *Psychological Research*, *81*(2), 432–444. https://doi.org/10.1007/s00426-016-0752-7

Ramachandran, V. S., & Anstis, S. M. (1983). Extrapolation of motion path in human visual perception. *Vision Research*, *23*(1), 83–85. https://doi.org/10.1016/0042-6989(83)90044-5

Rauch, N., Kaussner, A., Krüger, H.-P., Boverie, S., & Flemisch, F. (2009). The importance of driver state assessment within assessment within highly automated vehicles. *16th ITS World Congress and Exhibition on Intelligent Transport Systems and Services*, 1–8. http://www.haveit-eu.org/LH2Uploads/ItemsContent/25/3117-FULL-PAPER-THE-IMPORTANCE.pdf

Rawle, C. J., Chris Miall, R., & Praamstra, P. (2012). Fronto parietal theta activity supports behavioral decisions in movement-target selection. *Frontiers in Human Neuroscience*, *6*, 138. https://doi.org/10.3389/fnhum.2012.00138

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Reaz, M. B. I., Hussain, M. S., & Mohd-Yasin, F. (2006). Techniques of EMG signal analysis: detection, processing, classification and applications. *Biological Procedures Online*, *8*(1), 11–35. https://doi.org/10.1251/bpo115

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

Reeck, C., & Egner, T. (2014). Emotional task management: Neural correlates of switching between affective and non-affective task-sets. *Social Cognitive and Affective Neuroscience*, *10*(8), 1045–1053. https://doi.org/10.1093/scan/nsu153

Reid, G. B., & Nygren, T. E. (1988). The Subjective Workload Assessment Technique: A Scaling Procedure for Measuring Mental Workload. *Advances in Psychology*, *52*, 185–218. https://doi.org/10.1016/S0166-4115(08)62387-0

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*(1), 31–40. https://doi.org/10.1016/0028-3932(87)90041-8

Roach, B. J., & Mathalon, D. H. (2008). Event-related EEG time-frequency analysis: An overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophrenia Bulletin*, *34*(5), 907–926. https://doi.org/10.1093/schbul/sbn093

Robertshaw, K. D., & Wilkie, R. M. (2008). Does gaze influence steering around a bend? *Journal of Vision*, *8*(4), 1–13. https://doi.org/10.1167/8.4.18

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). “Oops!”: Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*(6), 747–758. https://doi.org/10.1016/S0028-3932(97)00015-8

Robertson, I. H., & O’Connell, R. (2012). Vigilant attention. In *A. C. Nobre & J. T. Coull (Eds.), Attention and Time* (pp. 79–88). New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199563456.003.0006

Rogers, R. D., & Monsell, S. (1995). Costs of a Predictable Switch Between Simple Cognitive Tasks. *Journal of Experimental Psychology: General*, *124*(2), 207–231. http://content.ebscohost.com/ContentServer.asp?T=P&P=AN&K=1995-31890-001&S=L&D=pdh&EbscoContent=dGJyMMTo50SeqLQ4zdnyOLCmr1Cep65Ssaq4TbeWxWXS&ContentCustomer=dGJyMPGnr0qxqK5QuePfgeyx44Dt6fIA

Rohenkohl, G., & Nobre, A. C. (2011). Alpha oscillations related to anticipatory attention follow temporal expectations. *Journal of Neuroscience*, *31*(40), 14076–14084. https://doi.org/10.1523/JNEUROSCI.3387-11.2011

Roman-Liu, D. (2016). The influence of confounding factors on the relationship between muscle contraction level and MF and MPF values of EMG signal: A review. *International Journal of Occupational Safety and Ergonomics*, *22*(1), 77–91. https://doi.org/10.1080/10803548.2015.1116817

Roman-Liu, D., Grabarek, I., Bartuzi, P., & Choromański, W. (2013). The influence of mental load on muscle tension. *Ergonomics*, *56*(7), 1125–1133. https://doi.org/10.1080/00140139.2013.798429

Rubino, C., Luksyte, A., Perry, S. J., & Volpone, S. D. (2009). How Do Stressors Lead to Burnout? The Mediating Role of Motivation. *Journal of Occupational Health Psychology*, *14*(3), 289–304. https://doi.org/10.1037/a0015284

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

Ruscio, D., Bos, A. J., & Ciceri, M. R. (2017). Distraction or cognitive overload? Using modulations of the autonomic nervous system to discriminate the possible negative effects of advanced assistance system. *Accident Analysis and Prevention*, *103*, 105–111. https://doi.org/10.1016/j.aap.2017.03.023

Rushworth, M. F. S., Krams, M., & Passingham, R. E. (2001). The attentional role of the left parietal cortex: The distinct lateralization and localization of motor attention in the human brain. *Journal of Cognitive Neuroscience*, *13*(5), 698–710. https://doi.org/10.1162/089892901750363244

Ruusuvuori, J. E., Aaltonen, T., Koskela, I., Ranta, J., Lonka, E., Salmenlinna, I., & Laakso, M. (2019). Studies on stigma regarding hearing impairment and hearing aid use among adults of working age: a scoping review. *Disability and Rehabilitation*, 1–11. https://doi.org/10.1080/09638288.2019.1622798

SAE (2018). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. Available Online at: Https://Www.Sae.Org/Standards/Content/J3016_201806/ (Accessed May 22 2020).

Saffarian, M., De Winter, J. C. F., & Happee, R. (2012). Automated driving: Human-factors issues and design solutions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2296–2300. https://doi.org/10.1177/1071181312561483

Sakihara, K., Hirata, M., Ebe, K., Kimura, K., Ryu, S. Y., Kono, Y., Muto, N., Yoshioka, M., Yoshimine, T., & Yorifuji, S. (2014). Cerebral oscillatory activity during simulated driving using MEG. *Frontiers in Human Neuroscience*, *8*, 975. https://doi.org/10.3389/fnhum.2014.00975

Salah, J., Abdelrahman, Y., Abdrabou, Y., Kassem, K., & Abdennadher, S. (2018). Exploring the usage of commercial bio-sensors for multitasking detection. *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*, 265–277. https://doi.org/10.1145/3282894.3282900

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, *39*(3), 490–500. https://doi.org/10.1016/j.ergon.2008.10.010

Sammer, G., Blecker, C., Gebhardt, H., Bischoff, M., Stark, R., Morgen, K., & Vaitl, D. (2007). Relationship between regional hemodynamic activity and simultaneously recorded EEG-theta associated with mental arithmetic-induced workload. *Human Brain Mapping*, *28*(8), 793–803. https://doi.org/10.1002/hbm.20309

Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: Where top-down meets bottom-up. *Brain Research Reviews*, *35*(2), 146–160. https://doi.org/10.1016/S0165-0173(01)00044-3

Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1015–1022. https://doi.org/10.1016/j.neubiorev.2009.12.006

Sauseng, P., Klimesch, W., Freunberger, R., Pecherstorfer, T., Hanslmayr, S., & Doppelmayr, M. (2006). Relevance of EEG alpha and theta oscillations during task switching. *Experimental Brain Research*, *170*(3), 295–301. https://doi.org/10.1007/s00221-005-0211-y

Savage, S. W., Potter, D. D., & Tatler, B. W. (2013). Does preoccupation impair hazard perception? A simultaneous EEG and Eye Tracking study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *17*, 52–62. https://doi.org/10.1016/j.trf.2012.10.002

Schewe, F., Cheng, H., Hafner, A., Sester, M., & Vollrath, M. (2019). Occupant Monitoring in Automated Vehicles: Classification of Situation Awareness Based on Head Movements While Cornering. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2078–2082. https://doi.org/10.1177/1071181319631048

Schier, M. A. (2000). Changes in EEG alpha power during simulated driving: A demonstration. *International Journal of Psychophysiology*, *37*(2), 155–162. https://doi.org/10.1016/S0167-8760(00)00079-9

Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accident Analysis and Prevention*, *41*(5), 1087–1093. https://doi.org/10.1016/j.aap.2009.06.007

Schmitz, C., Drake, L., Laake, M., Yin, P., & Pradarelli, R. (2012). Physiological Response to Fear in Expected and Unexpected Situations on Heart Rate, Respiration Rate and Horizontal Eye Movements. *Journal of Advanced Student Science*, *1*.

Schneegass, S., Pfleging, B., Broy, N., Schmidt, A., & Heinrich, F. (2013). A data set of real world driving to assess driver workload. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 150–157. https://doi.org/10.1145/2516540.2516561

Schneider, D. W. (2018). Alertness and cognitive control: Toward a spatial grouping hypothesis. *Attention, Perception, and Psychophysics*, *80*(4), 913–928. https://doi.org/10.3758/s13414-018-1491-1

Selcon, S. J., & M. Taylor, R. (1990). Evaluation of the Situational Awareness Rating Technique (SART) as a tool for aircrew systems design. *Journal of Aerospace Operations*, 62–66.

Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen’s f2, a measure of local effect size, from PROC MIXED. *Frontiers in Psychology*, *3*, 1–6. https://doi.org/10.3389/fpsyg.2012.00111

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

Shaffer, F., McCraty, R., & Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart’s anatomy and heart rate variability. *Frontiers in Psychology*, *5*, 1040. https://doi.org/10.3389/fpsyg.2014.01040

Shinohara, Y., Currano, R., Ju, W., & Nishizaki, Y. (2017). Visual attention during simulated autonomous driving in the US and Japan. *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 144–153. https://doi.org/10.1145/3122986.3122991

Simonet, M., Meziane, H. B., Runswick, O. R., North, J. S., Williams, A. M., Barral, J., & Roca, A. (2019). The modulation of event-related alpha rhythm during the time course of anticipation. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-54763-1

Singh, K. D., & Fawcett, I. P. (2008). Transient and linearly graded deactivation of the human default-mode network by a visual detection task. *NeuroImage*, *41*(1), 100–112. https://doi.org/10.1016/j.neuroimage.2008.01.051

Singmann, H., & Kellen, D. (2019). An Introduction to Mixed Models for Experimental Psychology. In *D. H. Spieler & E. Schumacher (Eds.), New Methods in Cognitive Psychology* (pp. 4–31). Hove: Psychology Press. https://doi.org/10.4324/9780429318405-2

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, *70*, 263–286. https://doi.org/10.1016/j.jbusres.2016.08.001

Smit, D. J. A., Posthuma, D., Boomsma, D. I., & De Geus, E. J. C. (2005). Heritability of background EEG across the power spectrum. *Psychophysiology*, *42*(6), 691–697.

https://doi.org/10.1111/j.1469-8986.2005.00352.x

Smith, S. S., Horswill, M. S., Chambers, B., & Wetton, M. (2009). Hazard perception in novice and experienced drivers: The effects of sleepiness. *Accident Analysis and Prevention*, *41*(4), 729–733. https://doi.org/10.1016/j.aap.2009.03.016

Sohn, M. H., Ursu, S., Anderson, J. R., Stenger, V. A., & Carter, C. S. (2000). The role of prefrontal cortex and posterior parietal cortex in task switching. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(24), 13448–13453. https://doi.org/10.1073/pnas.240460497

Solano Galvis, C. A., Tornay Mejías, F., & Gómez Milán, E. (2010). Effect of arousal increase in predictable and random task switching: evidence for the involvement of the anterior attentional network in random but not in predictable task switching. *Psicothema*, *22*(4), 703–707.

Solís-marcos, I., Galvao-carmona, A., & Kircher, K. (2017). Reduced attention allocation during short periods of partially automated driving: an event-related potentials study. *Frontiers in Human Neuroscience*, *11*, 1–13. https://doi.org/10.3389/fnhum.2017.00537

Sosnowski, T., Krzywosz-Rynkiewicz, B., & Roguska, J. (2004). Program running versus problem solving: Mental task effect on tonic heart rate. *Psychophysiology*, *41*(3), 467–475. https://doi.org/10.1111/j.1469-8986.2004.00171.x

Stadler, S., Cornet, H., Novaes Theoto, T., & Frenkler, F. (2019). A Tool, not a Toy: Using Virtual Reality to Evaluate the Communication Between Autonomous Vehicles and Pedestrians. *Augmented Reality and Virtual Reality*, 203–216. https://doi.org/10.1007/978-3-030-06246-0_15

Stanley, J., Peake, J. M., & Buchheit, M. (2013). Cardiac parasympathetic reactivation following exercise: Implications for training prescription. *Sports Medicine*, *43*(12), 1259–1277. https://doi.org/10.1007/s40279-013-0083-4

Stapel, J., Mullakkal-Babu, F. A., & Happee, R. (2019). Automated driving reduces perceived workload, but monitoring causes higher cognitive load than manual driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*, 590–605. https://doi.org/10.1016/j.trf.2018.11.006

Stein, P. K., & Pu, Y. (2012). Heart rate variability, sleep and sleep disorders. *Sleep Medicine Reviews*, *16*, 47–66. https://doi.org/10.1016/j.smrv.2011.02.005

Stothart, C., Mitchum, A., & Yehnert, C. (2015). The attentional cost of receiving a cell phone notification. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(4), 893–897. https://doi.org/10.1037/xhp0000100

Strauch, C., Mühl, K., Patro, K., Grabmaier, C., Reithinger, S., Baumann, M., & Huckauf, A. (2019). Real autonomous driving from a passenger's perspective: Two experimental investigations using gaze behaviour and trust ratings in field and simulator. *Transportation Research Part F: Traffic Psychology and Behaviour*, *66*, 15–28. https://doi.org/10.1016/j.trf.2019.08.013

Strayer, D. L., & Drews, F. A. (2007). Cell-Phone – Induced Driver Distraction. *Current Directions in Psychological Science*, *16*(3), 128–131.

Strobach, T., & Schubert, T. (2017). Mechanisms of Practice-Related Reductions of Dual-Task Interference with Simple Tasks: Data and Theory. *Advances in Cognitive Psychology*, *13*(1), 28–41. https://doi.org/10.5709/acp-0204-7

Sturm, W., & Willmes, K. (2001). On the functional neuroanatomy of intrinsic and phasic alertness. *NeuroImage*, *14*(1), S76–S84. https://doi.org/10.1006/nimg.2001.0839

Stuss, D. T., Shallice, T., Alexandar, M. P., & Picton, T. W. (1995). A Multidisciplinary Approach to Anterior Attentional Functions. *Annals of the New York Academy of Sciences*, *769*, 191–211. https://doi.org/10.1111/j.1749-6632.1995.tb38140.x

Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Industrial Psychiatry Journal*, *18*(1), 70–73. https://doi.org/10.4103/0972-6748.57865

Szabadi, E. (2015). Neuronal networks regulating sleep and arousal: Effect of drugs. In *Milestones in Drug Therapy*. Cham: Springer. https://doi.org/10.1007/978-3-319-11514-6_2

Taib, M. F. M., Bahn, S., & Yun, M. H. (2016). The effect of psychosocial stress on muscle activity during computer work: Comparative study between desktop computer and mobile computing products. *Work*, *54*(3), 543–555. https://doi.org/10.3233/WOR-162334

Takeda, Y., Sato, T., Kimura, K., Komine, H., Akamatsu, M., & Sato, J. (2016). Electrophysiological evaluation of attention in drivers and passengers: Toward an understanding of drivers' attentional state in autonomous vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, *42*, 140–150. https://doi.org/10.1016/j.trf.2016.07.008

Tang, Y. Y., Hölzel, B. K., & Posner, M. I. (2015). The neuroscience of mindfulness meditation. *Nature Reviews Neuroscience*, *16*, 213–225. https://doi.org/10.1038/nrn3916

Tang, Y. Y., Ma, Y., Wang, J., Fan, Y., Feng, S., Lu, Q., Yu, Q., Sui, D., Rothbart, M. K., Fan, M., & Posner, M. I. (2007). Short-term meditation training improves attention and self-regulation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(43), 17152–17156. https://doi.org/10.1073/pnas.0707678104

Tang, Y. Y., Yang, L., Leve, L. D., & Harold, G. T. (2012). Improving Executive Function and Its Neurobiological Mechanisms Through a Mindfulness-Based Intervention: Advances Within the Field of Developmental Neuroscience. *Child Development Perspectives*, *6*(4), 361–366.

https://doi.org/10.1111/j.1750-8606.2012.00250.x

Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology. (1996). Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, *17*, 354–381.

Tay, R., & Knowles, D. (2004). Drivers ’ Perception of Risks and Compensating Behaviours. *International Association of Traffic and Safety Sciences*, *28*(1), 89–94. https://doi.org/10.1016/S0386-1112(14)60095-9

Taylor-Phillips, S., Elze, M. C., Krupinski, E. A., Dennick, K., Gale, A. G., Clarke, A., & Mello-Thoms, C. (2014). Retrospective Review of the Drop in Observer Detection Performance Over Time in Lesion-enriched Experimental Studies. *Journal of Digital Imaging*, *28*(1), 32–40. https://doi.org/10.1007/s10278-014-9717-9

Thayer, J. F., Hansen, A. L., Saus-Rose, E., & Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: The neurovisceral integration perspective on self-regulation, adaptation, and health. *Annals of Behavioral Medicine*, *37*(2), 141–153. https://doi.org/10.1007/s12160-009-9101-z

Thayer, J. F., & Lane, R. D. (2000). A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of Affective Disorders*, *61*(3), 201–216. https://doi.org/10.1016/S0165-0327(00)00338-4

Thayer, J. F., Yamamoto, S. S., & Brosschot, J. F. (2010). The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. *International Journal of Cardiology*, *141*(2), 122–131. https://doi.org/10.1016/j.ijcard.2009.09.543

Thomson, D. R., Besner, D., & Smilek, D. (2015). A Resource-Control Account of Sustained Attention: Evidence From Mind-Wandering and Vigilance Paradigms. *Perspectives on Psychological Science*, *10*(1), 82–96. https://doi.org/10.1177/1745691614556681

Tichon, J. G., Mavin, T., Wallis, G., Visser, T. A. W., & Riek, S. (2014). Using Pupillometry and Electromyography to Track Positive and Negative Affect During Flight Simulation. *Aviation Psychology and Applied Human Factors*, *4*, 23–32. https://doi.org/10.1027/2192-0923/a000052

Toffetti, A., Wilschut, E. S., Martens, M. H., Schieben, A., Rambaldini, A., Merat, N., & Flemisch, F. (2009). CityMobil: Human factor issues regarding highly automated vehicles on eLane. *Transportation Research Record*, *2110*(1), 1–8. https://doi.org/10.3141/2110-01

Töllner, T., Wang, Y., Makeig, S., Müller, H. J., Jung, T. P., & Gramann, K. (2017). Two independent frontal midline theta oscillations during conflict detection and adaptation in a Simon-type manual reaching task. *Journal of Neuroscience*, *37*(9), 2504–2515. https://doi.org/10.1523/JNEUROSCI.1752-16.2017

Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin*, *20*(1), 12–16. https://doi.org/10.1093/oxfordjournals.bmb.a070274

Tsal, Y., & Benoni, H. (2010a). Diluting the Burden of Load: Perceptual Load Effects Are Simply Dilution Effects. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1645–1656. https://doi.org/10.1037/a0018172

Tsal, Y., & Benoni, H. (2010b). Much Dilution Little Load in Lavie and Torralbo's (2010) Response: A Reply. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1665–1668. https://doi.org/10.1037/a0021907

Tullis, T., & Albert, B. (2013). Behavioral and Physiological Metrics. In *Measuring the User Experience* (pp. 163–186). San Francisco, CA: Morgan Kauffman Publishers Inc. https://doi.org/10.1016/b978-0-12-415781-1.00007-8

Ulahannan, A., Jennings, P., Oliveira, L., & Birrell, S. (2020). Designing an adaptive interface: Using eye tracking to classify how information usage changes over time in partially automated vehicles. *IEEE Access*, *8*, 16865–16875. https://doi.org/10.1109/ACCESS.2020.2966928

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

Underwood, G., Phelps, N., Wright, C., Van Loon, E., & Galpin, A. (2005). Eye fixation scanpaths of younger and older drivers in a hazard perception task. *Ophthalmic and Physiological Optics*, *25*(4), 346–356. https://doi.org/10.1111/j.1475-1313.2005.00290.x

Unsworth, N., & Engle, R. W. (2008). Speed and Accuracy of Accessing Information in Working Memory: An Individual Differences Investigation of Focus Switching. *Journal of Experimental Psychology: Learning Memory and Cognition*, *34*(3), 616. https://doi.org/10.1037/0278-7393.34.3.616

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective and Behavioral Neuroscience*, *16*, 601–615. https://doi.org/10.3758/s13415-016-0417-4

Unsworth, N., & Robison, M. K. (2018). Tracking arousal state and mind wandering with pupillometry. *Cognitive, Affective and Behavioral Neuroscience*, *18*(4), 638–664. https://doi.org/10.3758/s13415-018-0594-4

Van Der Heiden, R. M. A., Janssen, C. P., Donker, S. F., Hardeman, L. E. S., Mans, K., & Kenemans, J. L. (2018). Susceptibility to audio signals during autonomous driving. *PLoS ONE*, *13*(8), e0201963. https://doi.org/10.1371/journal.pone.0201963 van der Meer, A. L. H., & van der Weel, F. R. (Ruud. (2017). Only three fingers write, but the whole brain works: A high-density EEG study showing advantages of drawing over typing for learning. *Frontiers in Psychology*, *8*, 706. https://doi.org/10.3389/fpsyg.2017.00706

van Dooren, M., de Vries, J. J. G. G. J., & Janssen, J. H. (2012). Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology and Behavior*, *106*(2), 298–304. https://doi.org/10.1016/j.physbeh.2012.01.020

van Eekelen, A. P. J., Houtveen, J. H., & Kerkhof, G. A. (2004). Circadian Variation in Cardiac Autonomic Activity: Reactivity Measurements to Different Types of Stressors. *Chronobiology International*, *21*(1), 107–129. https://doi.org/10.1081/CBI-120027983

van Lier, H. G., Pieterse, M. E., Garde, A., Postel, M. G., de Haan, H. A., Vollenbroek-Hutten, M. M. R., Schraagen, J. M., & Noordzij, M. L. (2020). A standardized validity assessment protocol for physiological signals from wearable technology: Methodological underpinnings and an application to the E4 biosensor. *Behavior Research Methods*, *52*(2), 607–629. https://doi.org/10.3758/s13428-019-01263-9

Van Vleet, T. M., DeGutis, J. M., Merzenich, M. M., Simpson, G. V., Zomet, A., & Dabit, S. (2016). Targeting alertness to improve cognition in older adults: A preliminary report of benefits in executive function and skill acquisition. *Cortex*, *82*, 100–118. https://doi.org/10.1016/j.cortex.2016.05.015

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task Switching: Interplay of Reconfiguration and Interference Control. *Psychological Bulletin*, *136*(4), 601–626. https://doi.org/10.1037/a0019791

Vartanian, L. R., Pinkus, R. T., & Smyth, J. M. (2018). Experiences of weight stigma in everyday life: Implications for health motivation. *Stigma and Health*, *3*(2), 85–92. https://doi.org/10.1037/sah0000077

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Verleger, R. (2020). Effects of relevance and response frequency on P3b amplitudes: Review of findings and comparison of hypotheses about the process reflected by P3b. *Psychophysiology*, *57*, e13542. https://doi.org/10.1111/psyp.13542

Visnovcova, Z., Mestanik, M., Gala, M., Mestanikova, A., & Tonhajzerova, I. (2016). The complexity of electrodermal activity is altered in mental cognitive stressors. *Computers in Biology and Medicine*, *79*, 123–129. https://doi.org/10.1016/j.compbiomed.2016.10.014

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

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

Vossel, S., Weidner, R., Thiel, C. M., & Fink, G. R. (2009). What is “Odd” in Posner’s Location-cueing Paradigm? Neural Responses to Unexpected Location and Feature Changes Compared. *Journal of Cognitive Neuroscience*, *21*(1), 30–41. https://doi.org/10.1162/jocn.2009.21003

Waersted, M., & Westgaard, R. H. (1996). Attention-related muscle activity in different body regions during vdu work with minimal physical activity. *Ergonomics*, *39*(4), 661–676. https://doi.org/10.1080/00140139608964488

Walker, F., Wang, J., Martens, M. H., & Verwey, W. B. (2019). Gaze behaviour and electrodermal activity: Objective measures of drivers’ trust in automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, *64*, 401–412. https://doi.org/10.1016/j.trf.2019.05.021

Wang, Y., Jung, T., & Lin, C. (2018). Theta and Alpha Oscillations in Attentional Interaction during Distracted Driving. *Frontiers in Behavioral Neuroscience*, *12*, 3. https://doi.org/10.3389/fnbeh.2018.00003

Ward, N., Hussey, E. K., Cunningham, E. C., Paul, E. J., McWilliams, T., & Kramer, A. F. (2019). Building the multitasking brain: An integrated perspective on functional brain activation during task-switching and dual-tasking. *Neuropsychologia*, *132*, 107149. https://doi.org/10.1016/j.neuropsychologia.2019.107149

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, *50*(3), 433–441. https://doi.org/10.1518/001872008X312152

Wascher, E., Rasch, B., Sänger, J., Hoffmann, S., Schneider, D., Rinkenauer, G., Heuer, H., & Gutberlet, I. (2014). Frontal theta activity reflects distinct aspects of mental fatigue. *Biological Psychology*, *96*(1), 57–65. https://doi.org/10.1016/j.biopsycho.2013.11.010

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology*, *46*(4), 361–413. https://doi.org/10.1016/S0010-0285(02)00520-0

Waterink, W., & van Boxtel, A. (1994). Facial and jaw-elevator EMG activity in relation to changes in performance level during a sustained information processing task. *Biological Psychology*, *37*(3), 183–198. https://doi.org/10.1016/0301-0511(94)90001-9

Weißbecker-Klaus, X., Ullsperger, P., Freude, G., & Schapkin, S. A. (2017). Impaired error processing and semantic processing during multitasking. *Journal of Psychophysiology*, *31*(4), 167–178. https://doi.org/10.1027/0269-8803/a000178

Welford, A. T. (1952). The “psychological refractory period” and the timing of high speed performance: A review and a theory. *British Journal of Psychology. General Section*, *43*(1), 2–19. https://doi.org/10.1111/j.2044-8295.1952.tb00322.x

Wetherell, M. A., & Carter, K. (2014). The multitasking framework: The effects of increasing workload on acute psychobiological stress reactivity. *Stress and Health*, *30*(2), 103–109. https://doi.org/10.1002/smi.2496

Wetherell, M. A., Craw, O., Smith, K., & Smith, M. A. (2017). Psychobiological responses to critically evaluated multitasking. *Neurobiology of Stress*, *7*, 68–73. https://doi.org/10.1016/j.ynstr.2017.05.002

Wetherell, M. A., & Sidgreaves, M. C. (2005). Short communication: Secretory immunoglobulin-A reactivity following increases in workload intensity using the Defined Intensity Stressor Simulation (DISS). *Stress and Health*, *21*(2), 99–106. https://doi.org/10.1002/smi.1038

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

Wijsman, J., Grundlehner, B., Liu, H., Penders, J., & Hermens, H. (2013). Wearable physiological sensors reflect mental stress state in office-like situations. *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 600–605. https://doi.org/10.1109/ACII.2013.105

Wijsman, J., Grundlehner, B., Penders, J., & Hermens, H. (2010). Trapezius muscle EMG as predictor of mental stress. *ACM Transactions on Embedded Computing Systems*, *2*(4), 1–20. https://doi.org/10.1145/2485984.2485987

Wilke, K., Martin, A., Terstegen, L., & Biel, S. S. (2007). A short history of sweat gland biology. *International Journal of Cosmetic Science*, *29*(3), 169–179. https://doi.org/10.1111/j.1467-2494.2007.00387.x

Wilson, D. E., Muroi, M., & MacLeod, C. M. (2011). Dilution, not load, affects distractor processing. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 319–335. https://doi.org/10.1037/a0021433

Winkel, J., Mathiassen, S. E., & Hägg, G. M. (1995). Normalization of upper trapezius EMG amplitude in ergonomic studies. *Journal of Electromyography and Kinesiology*, *5*(4), 195–196. https://doi.org/10.1016/1050-6411(96)85581-7

Wiradhany, W., & Koerts, J. (2019). Everyday functioning-related cognitive correlates of media multitasking: a mini meta-analysis. *Media Psychology*, 1–28.

https://doi.org/10.1080/15213269.2019.1685393

Wixted, F., & O' Sullivan, L. (2018). Effect of attention demand on upper trapezius muscle activity – A moderated mediation model. *International Journal of Industrial Ergonomics*, *66*, 146–156. https://doi.org/10.1016/j.ergon.2018.03.001

Wixted, F., O'Riordan, C., & O'Sullivan, L. (2018). Inhibiting the physiological stress effects of a sustained attention task on shoulder muscle activity. *International Journal of Environmental Research and Public Health*, *15*, 115. https://doi.org/10.3390/ijerph15010115

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*(6), 495–501. https://doi.org/10.1038/nrn1411

Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *The Journal of Neuroscience*, *20*(6), 1–20. https://doi.org/10.1523/jneurosci.20-06-j0002.2000

World Health Organisation. (2018). *Global Status Report on Road Safety*. https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/

Wulvik, A. S., Dybvik, H., & Steinert, M. (2020). Investigating the relationship between mental state (workload and affect) and physiology in a control room setting (ship bridge simulator). *Cognition, Technology and Work*, *22*(1), 95–108. https://doi.org/10.1007/s10111-019-00553-8

Xianglong, S., Hu, Z., Shumin, F., & Zhenning, L. (2018). Bus drivers ' mood states and reaction abilities at high temperatures. *Transportation Research Part F: Psychology and Behaviour*, *59*, 436–444. https://doi.org/10.1016/j.trf.2018.09.022

Yamada, E., Ogata, K., Kishimoto, J., Tanaka, M., Urakawa, T., Yamasaki, T., & Tobimatsu, S. (2015). Neural substrates of species-dependent visual processing of faces: Use of morphed faces. *Physiological Reports*, *3*(5), e12387. https://doi.org/10.14814/phy2.12387

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*, 459–482. https://doi.org/10.1002/cne.920180503

Yeung, N., Nystrom, L. E., Aronson, J. A., & Cohen, J. D. (2006). Between-task competition and cognitive control in task switching. *Journal of Neuroscience*, *26*(5), 1429–1438. https://doi.org/10.1523/JNEUROSCI.3109-05.2006

Yordanova, J., Kolev, V., & Polich, J. (2001). P300 and alpha event-related desynchronization (ERD). *Psychophysiology*, *38*, 143–152. https://doi.org/10.1017/S0048577201990079

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

Zakrzewska, M. Z., & Brzezicka, A. (2014). Working memory capacity as a moderator of load-related frontal midline theta variability in Sternberg task. *Frontiers in Human Neuroscience*, *8*, 399. https://doi.org/10.3389/fnhum.2014.00399

Zanto, T. P., & Gazzaley, A. (2014). Attention and ageing. In *A. C. Nobre & S. Kastner (Eds.), The Oxford Handbook of Attention*. New York, NY: Oxford University Press. https://doi.org/10.1017/CBO9781107415324.004

Zeeb, K., Buchner, A., & Schrauf, M. (2015). What determines the take-over time? An integrated model approach of driver take-over after automated driving. *Accident Analysis and Prevention*, *78*, 212–221. https://doi.org/10.1016/j.aap.2015.02.023

Zeeb, K., Buchner, A., & Schrauf, M. (2016). Is take-over time all that matters? the impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident Analysis and Prevention*, *92*, 230–239. https://doi.org/10.1016/j.aap.2016.04.002

Zeeb, K., Härtel, M., Buchner, A., & Schrauf, M. (2017). Why is steering not the same as braking ? The impact of non-driving related tasks on lateral and longitudinal driver interventions during conditionally automated driving. *Transportation Research Part F: Psychology and Behaviour*, *50*, 65–79. https://doi.org/10.1016/j.trf.2017.07.008

Zgaljardic, D. J., Borod, J. C., Foldi, N. S., & Mattis, P. (2003). A Review of the Cognitive and Behavioral Sequelae of Parkinson's Disease: Relationship to Frontostriatal Circuitry. *Cognitive and Behavioral Neurology*, *16*(4), 193–210. https://doi.org/10.1097/00146965-200312000-00001

Zhang, B., de Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *64*, 285–307. https://doi.org/10.1016/j.trf.2019.04.020

Zhang, C., Sun, W., Song, Q., Gu, H., & Mao, D. (2018). Performance of older adults under dual task during stair descent. *Journal of Exercise Science and Fitness*, *16*(3), 99–105. https://doi.org/10.1016/j.jesf.2018.09.001

Zhao, C., Zhao, M., Liu, J., & Zheng, C. (2012). Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator. *Accident Analysis and Prevention*, *45*, 83–90. https://doi.org/10.1016/j.aap.2011.11.019

Zheng, R., Yamabe, S., Nakano, K., & Suda, Y. (2015). Biosignal analysis to assess mental stress in automatic driving of trucks: Palmar perspiration and masseter electromyography. *Sensors*, *15*(3), 5136–5150. https://doi.org/10.3390/s150305136

Zimmermann-Viehoff, F., Thayer, J., Koenig, J., Herrmann, C., Weber, C. S., & Deter, H. C. (2016). Short-term effects of espresso coffee on heart rate variability and blood pressure in habitual and non-habitual coffee consumers – A randomized crossover study. *Nutritional Neuroscience*, *19*(4), 169–175. https://doi.org/10.1179/1476830515Y.0000000018

# Appendices

Appendix 2.1. Table summarising raw means (SD) of subjective and physiological data at one-minute intervals at baseline, during task, and at recovery for low and high workload multitasking

| Measure | Workload | Minute periods over task | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Baseline | | Task | | | | | | | | | | Recovery | |
| | | 1 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 1 | 2 |
| Subjective (1-10) | Low | 1.51 | 1.49 | 2.91 | 3.37 | 3.83 | 4.00 | 4.07 | 4.16 | 4.46 | 4.44 | 4.44 | 4.60 | 2.40 | 1.74 |
| | | (1.34) | (1.26) | (1.44) | (1.64) | (1.68) | (1.67) | (1.9) | (1.75) | (1.76) | (1.93) | (1.96) | (1.93) | (1.64) | (1.03) |
| | High | 1.84 | 1.77 | 4.09 | 4.74 | 5.30 | 5.56 | 5.72 | 5.86 | 5.93 | 5.95 | 5.95 | 6.00 | 3.07 | 2.02 |
| | | (1.58) | (1.63) | (1.58) | (1.82) | (1.72) | (1.97) | (1.99) | (2.01) | (1.99) | (2.01) | (2.07) | (1.98) | (1.7) | (1.53) |
| Heart rate (bpm) | Low | 74.40 | 73.20 | 74.42 | 72.74 | 73.72 | 74.04 | 74.19 | 74.47 | 74.67 | 75.13 | 75.48 | 75.25 | 75.24 | 73.18 |
| | | (9.83) | (9.65) | (9.50) | (9.64) | (9.54) | (9.58) | (9.72) | (9.63) | (9.77) | (9.23) | (9.43) | (9.74) | (10.04) | (9.74) |
| | High | 73.78 | 73.16 | 74.81 | 73.46 | 73.93 | 74.09 | 74.23 | 74.58 | 74.47 | 74.91 | 74.44 | 74.63 | 74.77 | 72.51 |
| | | (10.28) | (10.10) | (9.39) | (9.59) | (9.57) | (9.63) | (8.89) | (8.54) | (8.76) | (9.22) | (9.35) | (9.12) | (10.12) | (9.21) |
| RMSSD (msecs) | Low | 64.41 | 74.37 | 54.04 | 57.79 | 55.58 | 51.82 | 69.45 | 51.49 | 70.48 | 53.93 | 62.02 | 59.00 | 69.89 | 54.39 |
| | | (45.79) | (110.28) | (34.62) | (38.22) | (37.7) | (34.48) | (100.21) | (37.7) | (103.67) | (35.42) | (56.85) | (55.9) | (68.85) | (34.68) |
| | High | 92.45 | 83.52 | 79.96 | 78.79 | 80.53 | 84.64 | 73.92 | 73.47 | 74.69 | 73.26 | 77.41 | 74.51 | 91.08 | 86.36 |
| | | (129.24) | (138.36) | (146.07) | (141.27) | (136.79) | (145.55) | (130.68) | (126.85) | (126.26) | (128.36) | (128.7) | (128.00) | (128.15) | (130.75) |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SDSD (msecs) | Low | 64.38 | 74.36 | 54.00 | 57.78 | 55.57 | 51.81 | 69.44 | 51.48 | 70.47 | 53.91 | 62.00 | 58.99 | 69.86 | 54.38 |
| | | (46.21) | (111.68) | (35.03) | (38.67) | (38.11) | (34.78) | (101.48) | (38.13) | (104.90) | (35.85) | (57.57) | (56.61) | (69.73) | (34.77) |
| | High | 92.44 | 83.51 | 79.93 | 78.77 | 80.52 | 84.63 | 73.91 | 73.45 | 74.68 | 73.24 | 77.39 | 74.49 | 91.02 | 87.76 |
| | | (130.81) | (140.07) | (147.93) | (143.05) | (138.51) | (147.39) | (132.34) | (128.46) | (127.87) | (129.98) | (130.33) | (129.63) | (129.77) | (133.71) |
| pNN50 (%) | Low | 28.55 | 28.01 | 27.11 | 30.35 | 28.67 | 26.73 | 27.80 | 27.10 | 27.10 | 26.82 | 26.58 | 26.76 | 26.43 | 27.05 |
| | | (21.54) | (21.05) | (20.23) | (23.58) | (23.94) | (22.71) | (23.63) | (20.89) | (23.01) | (21.11) | (23.24) | (23.51) | (19.80) | (20.81) |
| | High | 33.23 | 32.73 | 28.43 | 29.77 | 30.66 | 28.81 | 28.36 | 28.36 | 29.89 | 28.17 | 28.44 | 29.70 | 30.83 | 31.19 |
| | | (25.47) | (26.43) | (22.78) | (25.76) | (27.15) | (26.60) | (24.84) | (24.66) | 2(4.35) | (24.14) | (24.20) | (25.07) | 20.68) | (24.98) |
| NS-SCR freq | Low | 2.74 | 2.15 | 3.41 | 2.62 | 1.82 | 1.92 | 1.90 | 1.92 | 1.72 | 1.67 | 1.87 | 1.85 | 2.44 | 1.85 |
| | | (1.64) | (1.51) | (2.03) | (2.27) | (2.24) | (2.03) | (1.96) | (2.3) | (1.45) | (1.76) | (1.99) | (1.94) | (1.34) | (1.44) |
| | High | 2.23 | 1.85 | 4.00 | 2.82 | 2.41 | 2.23 | 2.05 | 1.82 | 1.92 | 1.87 | 1.77 | 1.77 | 2.31 | 1.90 |
| | | (1.67) | (1.59) | (2.20) | (2.04) | (1.71) | (1.91) | (1.74) | (1.32) | (1.25) | (1.56) | (1.37) | (1.42) | (1.57) | (1.48) |
| NS-SCR amp (µS) | Low | 8.07 | 7.49 | 9.41 | 7.28 | 5.77 | 5.94 | 6.25 | 6.27 | 6.44 | 5.35 | 5.62 | 5.56 | 8.16 | 7.00 |
| | | (4.02) | (3.97) | (3.93) | (4.61) | (4.95) | (4.47) | (4.52) | (4.71) | (4.44) | (4.66) | (4.58) | (4.79) | (3.66) | (4.47) |
| | High | 7.80 | 7.07 | 9.41 | 7.92 | 7.22 | 6.54 | 6.79 | 6.71 | 7.04 | 6.38 | 6.49 | 6.25 | 8.22 | 7.22 |
| | | (4.53) | (4.62) | (4.30) | (4.38) | (3.93) | (4.24) | (4.07) | (3.84) | (3.58) | (4.12) | (4.01) | (4.13) | (4.19) | (4.6) |

Appendix 2.2. Table summarising raw means (SD) of alpha and theta activity at baseline, during task, and at recovery for low and high workload multitasking

| Frequency (Hz) | Electrodes | Workload | Interval | | |
|---|---|---|---|---|---|
| | | | Baseline | Task | Recovery |
| Theta (4 – 7) | F3, F4, F7, F8, Fz | Low | 1.35 (0.37) | 1.52 (0.33) | 1.53 (0.37) |
| | | High | 1.24 (0.37) | 1.36 (0.33) | 1.36 (0.37) |
| Alpha (8.5 – 12.5) | F3, F4, F7, F8, Fz | Low | 1.35 (0.37) | 1.52 (0.35) | 1.53 (0.35) |
| | | High | 1.24 (0.37) | 1.36 (0.33) | 1.36 (0.37) |
| Theta (4 – 7) | P3, P4, P7, P8, Pz | Low | 1.33 (0.49) | 1.53 (0.44) | 1.61 (0.49) |
| | | High | 1.23 (0.49) | 1.37 (0.44) | 1.46 (0.49) |
| Alpha (8.5 – 12.5) | P3, P4, P7, P8, Pz | Low | 1.18 (0.42) | 1.25 (0.42) | 1.48 (0.42) |
| | | High | 1.07 (0.42) | 1.06 (0.42) | 1.34 (0.42) |

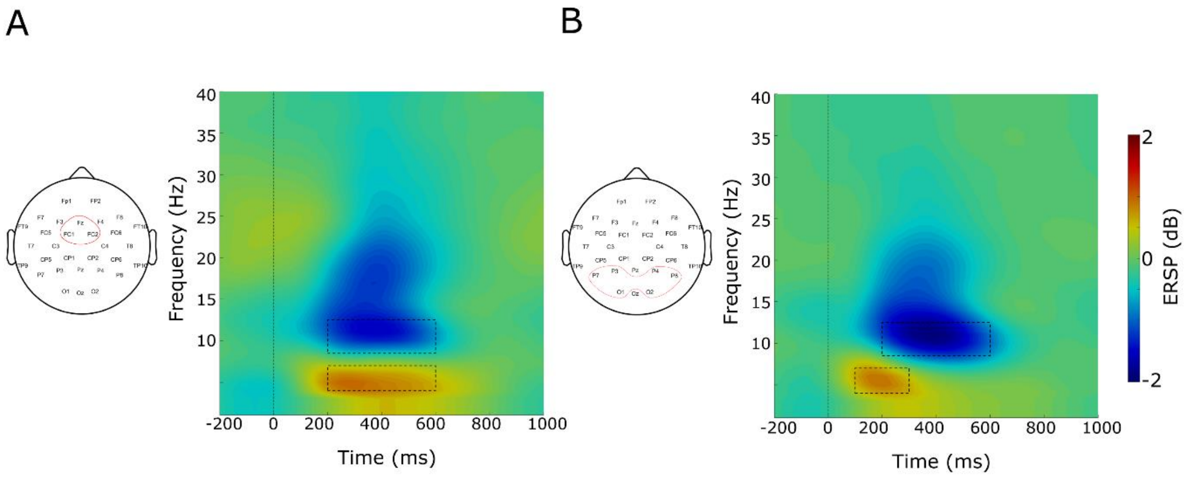
Appendix 2.3. Table summarising natural logarithm transformed means (SD) of alpha and theta activity at one-minute intervals at baseline, during task, and at recovery for low and high workload multitasking

| Frequency (Hz) | Electrodes | Workload | Minute periods over task | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Baseline | | Task | | | | | | | | | | Recovery | |
| | | | 1 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 1 | 2 |
| Theta | F3, F4, F7, | Low | 1.34 | 1.36 | 1.44 | 1.45 | 1.47 | 1.49 | 1.53 | 1.52 | 1.54 | 1.55 | 1.58 | 1.60 | 1.54 | 1.53 |
| (4 – 7) | F8, Fz | | (0.37) | (0.37) | (0.39) | (0.38) | (0.40) | (0.43) | (0.45) | (0.49) | (0.49) | (0.50) | (0.53) | (0.56) | (0.58) | (0.57) |
| | | High | 1.26 | 1.21 | 1.26 | 1.32 | 1.35 | 1.34 | 1.36 | 1.36 | 1.38 | 1.40 | 1.40 | 1.44 | 1.36 | 1.37 |
| | | | (0.29) | (0.27) | (0.30) | (0.29) | (0.30) | (0.31) | (0.32) | (0.33) | (0.35) | (0.35) | (0.36) | (0.38) | (0.42) | (0.41) |
| Alpha | | Low | 1.11 | 1.13 | 1.07 | 1.09 | 1.11 | 1.13 | 1.17 | 1.16 | 1.19 | 1.19 | 1.22 | 1.25 | 1.36 | 1.35 |
| (8.5 – 12.5) | | | (0.36) | (0.37) | (0.35) | (0.35) | (0.37) | (0.39) | (0.41) | (0.43) | (0.44) | (0.44) | (0.47) | (0.49) | (0.54) | (0.52) |
| | | High | 1.05 | 1.02 | 0.92 | 0.97 | 1.00 | 1.00 | 1.02 | 1.02 | 1.04 | 1.05 | 1.07 | 1.10 | 1.21 | 1.23 |
| | | | (0.31) | (0.3) | (0.28) | (0.27) | (0.28) | (0.30) | (0.29) | (0.30) | (0.31) | (0.34) | (0.34) | (0.38) | (0.43) | (0.42) |
| Theta | P3, P4, P7, | Low | 1.32 | 1.33 | 1.4 | 1.43 | 1.46 | 1.49 | 1.53 | 1.54 | 1.58 | 1.60 | 1.65 | 1.67 | 1.6 | 1.61 |
| (4 – 7) | P8, Pz | | (0.44) | (0.48) | (0.55) | (0.55) | (0.57) | (0.61) | (0.62) | (0.67) | (0.69) | (0.71) | (0.74) | (0.77) | (0.73) | (0.75) |
| | | High | 1.27 | 1.18 | 1.25 | 1.30 | 1.34 | 1.33 | 1.36 | 1.36 | 1.40 | 1.41 | 1.43 | 1.48 | 1.43 | 1.44 |
| | | | (0.4) | (0.34) | (0.38) | (0.38) | (0.39) | (0.38) | (0.41) | (0.38) | (0.43) | (0.44) | (0.45) | (0.49) | (0.48) | (0.47) |
| Alpha | | Low | 1.18 | 1.19 | 1.1 | 1.14 | 1.17 | 1.20 | 1.25 | 1.26 | 1.29 | 1.31 | 1.35 | 1.38 | 1.47 | 1.49 |
| (8.5 – 12.5) | | | (0.45) | (0.47) | (0.54) | (0.54) | (0.57) | (0.59) | (0.61) | (0.64) | (0.66) | (0.68) | (0.71) | (0.73) | (0.67) | (0.68) |
| | | High | 1.1 | 1.05 | 0.94 | 0.99 | 1.03 | 1.04 | 1.06 | 1.05 | 1.10 | 1.12 | 1.13 | 1.18 | 1.33 | 1.35 |
| | | | (0.36) | (0.34) | (0.31) | (0.31) | (0.31) | (0.32) | (0.33) | (0.33) | (0.35) | (0.37) | (0.38) | (0.42) | (0.45) | (0.44) |

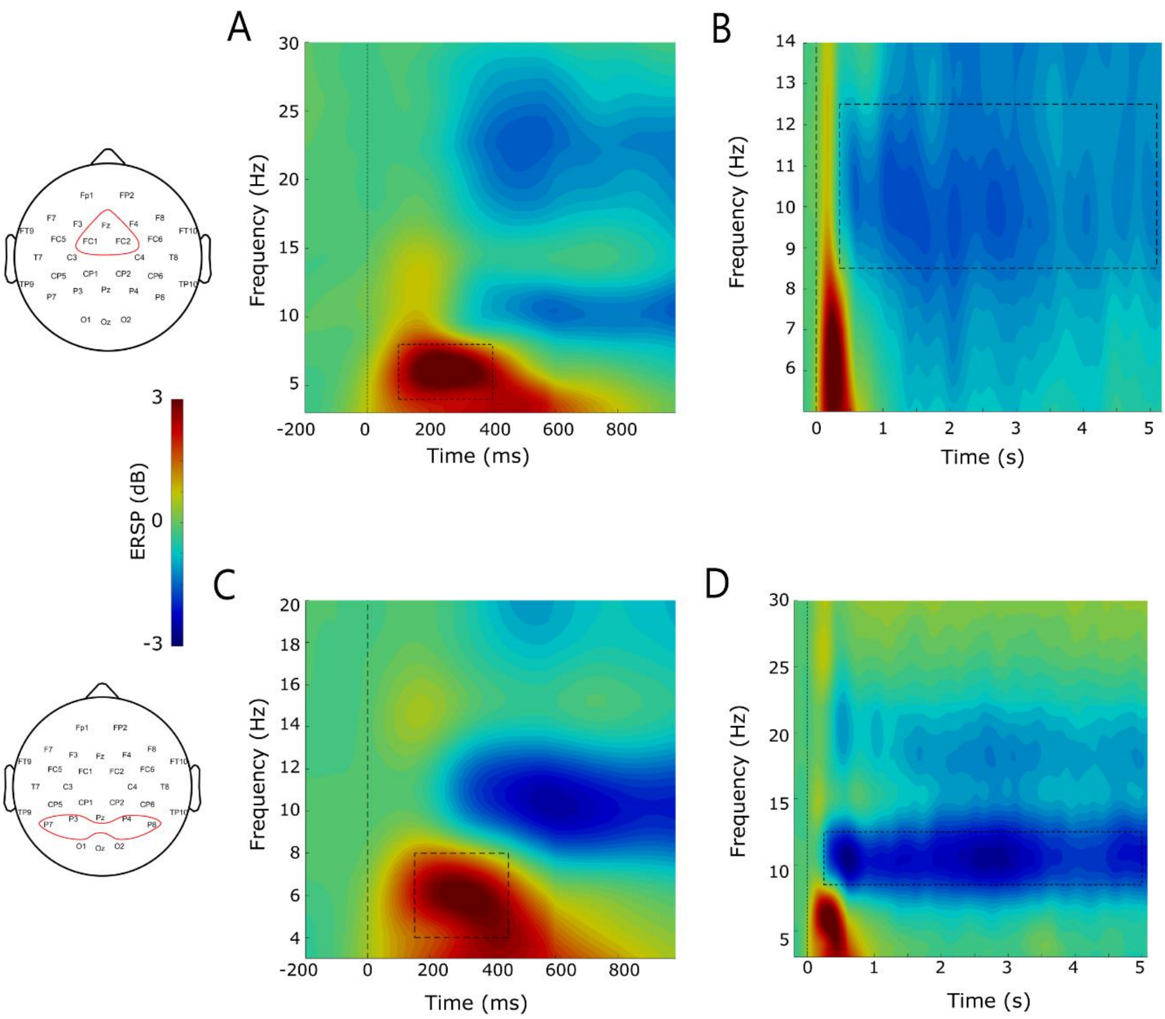
Appendix 2.4. Table summarising raw means (SD) of alpha and theta activity at baseline, during task, and at recovery for low and high workload multitasking

| Frequency (Hz) | Electrodes | Workload | Minute periods over task | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Baseline | | Task | | | | | | | | | | Recovery | |
| | | | 1 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 1 | 2 |
| Theta | F3, F4, F7, F8, | Low | 3.82 | 3.90 | 4.22 | 4.26 | 4.35 | 4.44 | 4.62 | 4.57 | 4.66 | 4.71 | 4.85 | 4.95 | 4.66 | 4.62 |
| (4 – 7) | Fz | | (1.45) | (1.45) | (1.48) | (1.46) | (1.49) | (1.49) | (1.57) | (1.60) | (1.63) | (1.65) | (1.70) | (1.75) | (1.79) | (1.77) |
| | | High | 3.53 | 3.35 | 3.53 | 3.74 | 3.86 | 3.82 | 3.90 | 3.90 | 3.97 | 4.06 | 4.06 | 4.22 | 3.90 | 3.94 |
| | | | (1.34) | (1.31) | (2.46) | (1.34) | (1.35) | (1.36) | (1.38) | (1.39) | (1.42) | (1.42) | (1.43) | (1.46) | (1.52) | (1.51) |
| Alpha | | Low | 3.03 | 3.10 | 2.92 | 2.97 | 3.03 | 3.10 | 3.22 | 3.19 | 3.29 | 3.29 | 3.39 | 3.49 | 3.90 | 3.86 |
| (8.5 – 12.5) | | | (1.43) | (1.45) | (1.42) | (1.42) | (1.45) | (1.48) | (1.51) | (1.54) | (1.55) | (1.55) | (1.60) | (1.63) | (1.72) | (1.68) |
| | | High | 2.86 | 2.77 | 2.51 | 2.64 | 3.60 | 3.67 | 2.77 | 2.77 | 2.83 | 2.86 | 2.92 | 3.00 | 3.35 | 3.42 |
| | | | (1.36) | (1.35) | (1.32) | (1.31) | (1.32) | (1.35) | (1.34) | (1.35) | (1.36) | (1.40) | (1.40) | (1.46) | (1.54) | (1.52) |
| Theta | P3, P4, P7, P8, | Low | 3.74 | 3.78 | 4.06 | 4.18 | 4.31 | 4.44 | 4.62 | 4.66 | 4.85 | 4.95 | 5.21 | 5.31 | 4.95 | 5.00 |
| (4 – 7) | Pz | | (1.55) | (1.62) | (1.73) | (1.73) | (1.77) | (1.84) | (1.86) | (1.95) | (1.99) | (2.03) | (2.10) | (2.16) | (2.08) | (2.12) |
| | | High | 3.56 | 3.25 | 3.49 | 3.67 | 3.82 | 3.78 | 3.90 | 3.90 | 4.06 | 4.10 | 4.18 | 4.39 | 4.18 | 4.22 |
| | | | (1.49) | (1.40) | (1.46) | (1.46) | (1.48) | (1.46) | (1.51) | (1.46) | (1.54) | (1.55) | (1.57) | (1.63) | (1.62) | (1.60) |
| Alpha | | Low | 3.25 | 3.29 | 3.00 | 3.13 | 3.22 | 3.32 | 3.49 | 3.53 | 3.63 | 3.71 | 3.86 | 3.97 | 4.35 | 4.44 |
| (8.5 – 12.5) | | | (1.57) | (1.60) | (1.72) | (1.72) | (1.77) | (1.80) | (1.84) | (1.90) | (1.93) | (1.97) | (2.03) | (2.08) | (1.95) | (1.97) |
| | | High | 3.00 | 2.86 | 2.56 | 2.69 | 2.80 | 2.83 | 2.89 | 2.86 | 3.00 | 3.06 | 3.10 | 3.25 | 3.78 | 3.86 |
| | | | (1.48) | (1.40) | (1.36) | (1.36) | (1.36) | (1.38) | (1.39) | (1.39) | (1.42) | (1.45) | (1.46) | (1.52) | (1.57) | (1.55) |

Appendix 3.1 Condition grand-averaged time-frequency spectrograms for visual search. Regions of interest were defined for alpha (8.5 - 12.5Hz) and theta (4 - 7Hz) power. **(A)** Frontal activity was averaged at electrodes Fz, FC1, FC2 between time 200 - 600ms for alpha, and 200 - 600ms for theta. **(B)** Parietal-occipital activity was averaged at electrodes O1, O2, P3, P4, P7, P8 between time 200 - 600ms for alpha, and 100 - 300ms for theta.



Appendix 3.2. Condition grand-averaged time-frequency spectrograms for 5 second tracking task. **(A)** Frontal activity was averaged at electrodes Fz, FC1, FC2 between 100 - 400ms for theta (4-8Hz) activity and **(B)** 350 - 5100ms for alpha activity (8.5 – 12.5 Hz). **(C)** Parietal activity was averaged over contralateral electrodes (P3, P7) and ipsilateral electrodes (P4, P8) between 150 – 450ms for theta activity (4 – 8 Hz) and **(D)** 250 – 5100ms for alpha (8.5 – 12.5 Hz) activity.



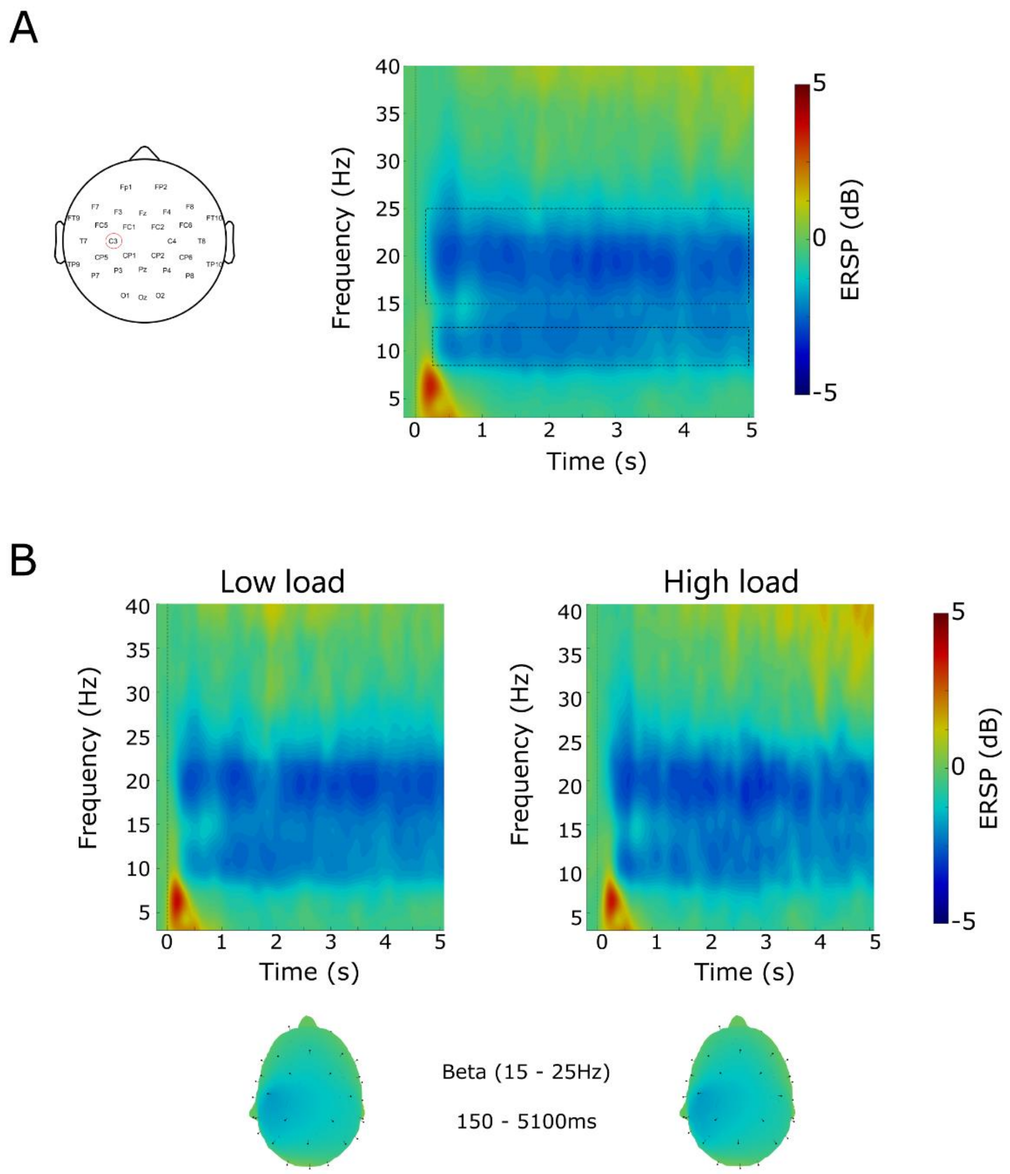
Appendix 3.3 Table representing significant results for paired-samples *t*-test contrasting low and high load over evoked frontal theta power (dB normalised) during the 5 second tracking task. Higher scores represent greater theta synchronisation.

| Time point | Time period (ms) | Low load | High load | *t*-statistic | *p*-value |
|---|---|---|---|---|---|
| 1 | 110 | 3.21 | 1.57 | 10.54 | < .001 |
| 2 | 142 | 3.49 | 2.23 | 7.47 | < .001 |
| 3 | 172 | 3.55 | 2.82 | 4.04 | < .001 |
| 6 | 268 | 2.94 | 3.46 | -2.54 | .02 |
| 7 | 300 | 2.64 | 3.31 | 3.40 | < .001 |
| 8 | 330 | 2.35 | 3.07 | -3.89 | < .001 |

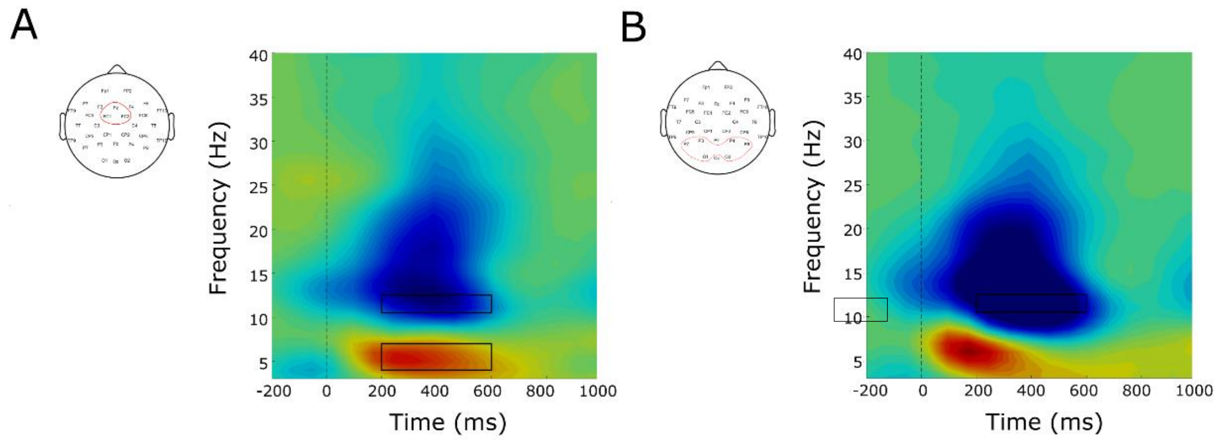
Appendix 3.4. Table representing significant results for paired-samples *t*-test contrasting low and high load over contralateral alpha desynchronisation (dB normalised) during the 5 second tracking task. Lower scores represent greater alpha desynchronisation.

| Time point | Time period (ms) | Low load | High load | *t*-statistic | *p*-value |
|---|---|---|---|---|---|
| 2 | 268 | 0.99 | 0.56 | -3.62 | .004 |
| 11 | 552 | -1.81 | -2.27 | 2.82 | .008 |
| 12 | 584 | -1.77 | -2.31 | 3.31 | .002 |
| 14 | 646 | -1.67 | -2.24 | 3.57 | .001 |
| 15 | 678 | -1.63 | -2.16 | 3.29 | .002 |
| 16 | 710 | -1.62 | -2.08 | 2.84 | .008 |
| 116 | 3868 | -1.69 | -2.09 | 2.68 | .01 |
| 117 | 3898 | -1.69 | -2.08 | 2.79 | .008 |
| 118 | 3930 | -1.65 | -2.06 | 2.94 | .006 |
| 119 | 3962 | -1.62 | -2.04 | 3.13 | .004 |
| 120 | 3994 | -1.62 | -2.04 | 3.30 | .002 |
| 121 | 4026 | -1.62 | -2.04 | 3.33 | .002 |
| 22 | 4056 | -1.64 | -2.03 | 3.08 | .004 |
| 130 | 4310 | -1.61 | -2.04 | 2.76 | .009 |
| 131 | 4342 | -1.62 | -2.10 | 3.31 | .002 |
| 132 | 4372 | -1.64 | -2.15 | 3.50 | .001 |
| 133 | 4404 | -1.68 | -2.16 | 3.29 | .002 |
| 134 | 4436 | -1.71 | -2.14 | 2.92 | .006 |
| 139 | 4594 | -1.81 | -2.17 | 2.74 | .01 |
| 140 | 4626 | -1.80 | -2.19 | 2.83 | .008 |
| 141 | 4656 | -1.79 | -2.20 | 2.88 | .007 |
| 142 | 4688 | -1.77 | -2.21 | 2.95 | .006 |
| 143 | 4720 | -1.76 | -2.22 | 3.04 | .005 |
| 144 | 4752 | -1.77 | -2.23 | 3.10 | .004 |
| 145 | 4784 | -1.80 | -2.24 | 3.11 | .004 |
| 146 | 4814 | -1.82 | -2.26 | 3.08 | .004 |
| 147 | 4846 | -1.83 | -2.27 | 3.09 | .004 |
| 148 | 4878 | -1.81 | -2.28 | 3.23 | .003 |
| 149 | 4910 | -1.78 | -2.28 | 3.42 | .002 |
| 150 | 4942 | -1.75 | -2.27 | 3.56 | .001 |
| 152 | 5004 | -1.71 | -2.21 | 3.53 | .001 |
| 153 | 5036 | -1.70 | -2.18 | 3.45 | .002 |
| 154 | 5068 | -1.69 | -2.16 | 3.43 | .002 |
| 155 | 5098 | -1.68 | -2.16 | 3.41 | .002 |

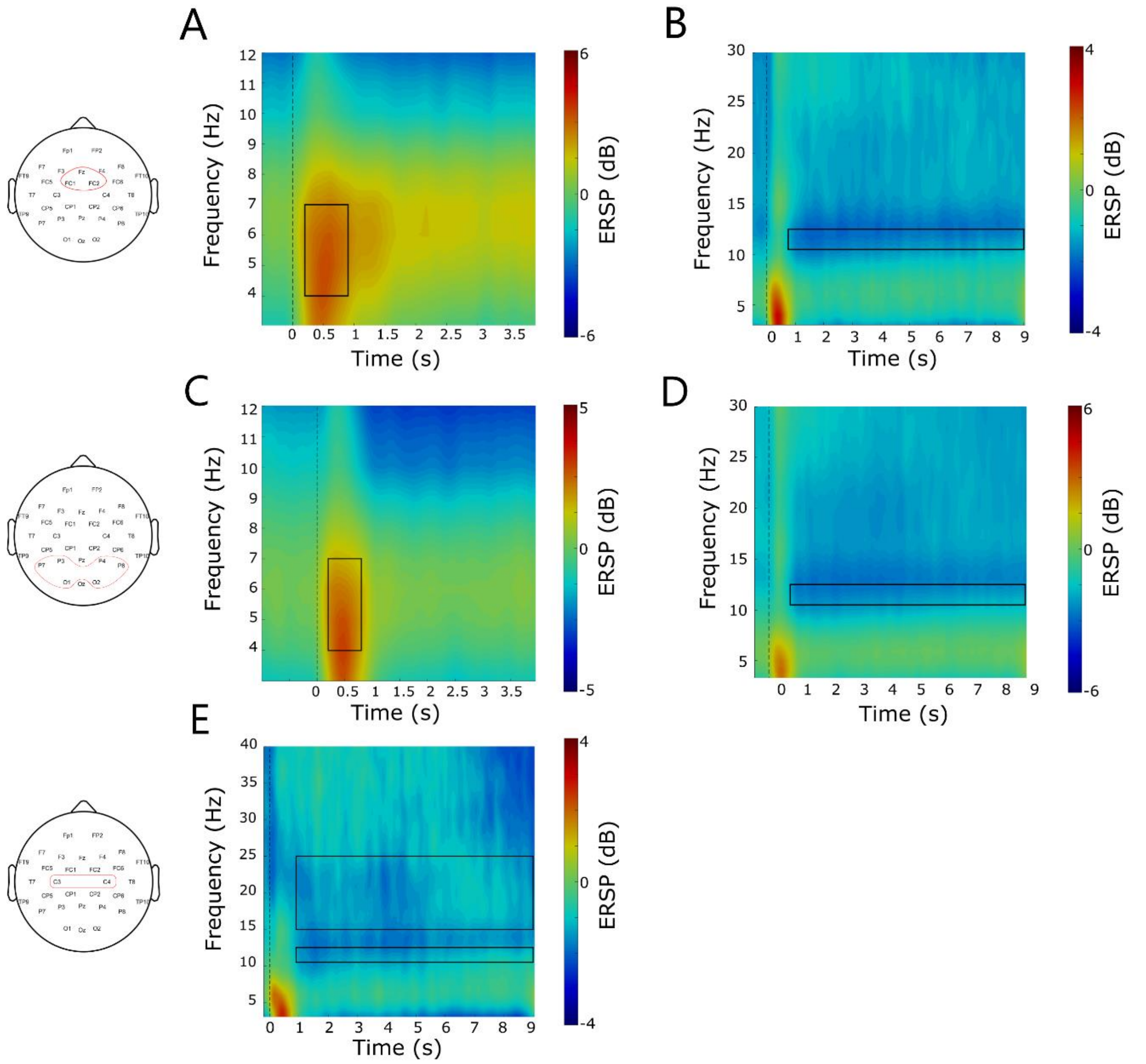
Appendix 3.5. Motor-related oscillations during the tracking task. **(A)** Condition (low and high load) grand-averaged time-frequency spectrogram for tracking task. Motor activity was extracted from electrode C3 between time 150 - 5100ms for beta (15- 25 Hz) activity; and between time points 250 – 5100 ms for alpha activity (8.5 – 12.5 Hz). **(B)** Grand-averaged time-frequency spectrograms for motor (C3) tracking task activity. Averaged topographical plots (150 – 1500 ms) over broadband beta (15 – 25 Hz) activity.



Appendix 4.1. Condition grand-averaged time-frequency spectrograms for visual search. Regions of interest were defined for alpha (10.5 - 12.5Hz) and theta (4 - 7Hz) power. **(A)** Frontal activity was averaged at electrodes Fz, FC1, FC2 between time 200 - 600ms for alpha, and 200 - 600ms for theta. **(B)** Parieto-occipital activity was averaged at electrodes O1, O2, P3, P4, P7, P8 between time 200 - 600ms for alpha and 100 – 300 ms for theta.



Appendix 4.2. Condition grand-averaged time-frequency spectrograms for simulated driving. **(A)** Frontal activity was averaged at electrodes Fz, FC1, FC2 between time 200 – 900 ms for theta (4 – 7 Hz) activity. **(B)** Frontal activity was averaged at electrodes Fz, FC1, FC2 between time 900 – 9100 ms for alpha (10.5 – 12.5 Hz) activity. **(C)** Parietal theta (4 – 7 Hz) activity was averaged at over parieto-occipital electrodes (P3, P7, P4, P8, O1, O2) between time 200 – 800 ms. **(D)** Parietal upper alpha (10.5 - 12.5 Hz) activity was averaged at over parietal-occipital electrodes (P3, P7, P4, P8, O1, O2) between time 900 – 9100 ms. **(E)** Motor-related activity was averaged over motor cortex (C3, C4) between 950 – 9100 ms.



Appendix 4.3. Table representing significant results for paired-samples *t*-test over parieto-occipital alpha desynchronisation between passive viewing and manual driving.

| Time point | Time period (ms) | Task | | *t*-statistic | *p*-value |
|---|---|---|---|---|---|
| | | Manual driving | Passive viewing | | |
| 1 | 998 | -2.72 | -1.20 | -4.53 | < .001 |
| 2 | 1072 | -2.76 | -1.27 | -4.30 | < .001 |
| 3 | 1148 | -2.72 | -1.31 | -3.80 | < .001 |
| 4 | 1222 | -2.69 | -1.35 | -3.59 | < .001 |
| 5 | 1296 | -2.61 | -1.46 | -3.45 | .002 |
| 6 | 1372 | -2.65 | -1.48 | -3.46 | .002 |
| 7 | 1446 | -2.71 | -1.40 | -3.45 | .002 |
| 8 | 1522 | -2.79 | -1.29 | -3.51 | < .001 |
| 9 | 1596 | -2.90 | -1.18 | -3.98 | < .001 |
| 10 | 1670 | -2.88 | -1.18 | -4.53 | < .001 |
| 11 | 1746 | -2.82 | -1.22 | -4.66 | < .001 |
| 12 | 1820 | -2.79 | -1.32 | -4.27 | < .001 |
| 13 | 1896 | -2.81 | 1.46 | -3.78 | < .001 |
| 14 | 1970 | -2.85 | -1.52 | -3.66 | < .001 |
| 15 | 2044 | -2.82 | -1.47 | -3.96 | < .001 |
| 16 | 2120 | -2.80 | -1.36 | -4.10 | < .001 |
| 17 | 2194 | -2.80 | -1.28 | -3.92 | < .001 |
| 18 | 2270 | -2.89 | -1.32 | -3.93 | < .001 |
| 19 | 2344 | -2.93 | -1.48 | -3.52 | < .001 |
| 20 | 2418 | -2.94 | -1.59 | -2.99 | .005 |
| 21 | 2494 | -2.96 | -1.61 | -2.93 | .006 |
| 22 | 2568 | -2.95 | -1.63 | -3.02 | .005 |
| 23 | 2642 | -2.87 | -1.70 | -2.98 | .006 |
| 24 | 2718 | -2.85 | -1.76 | -2.98 | .003 |
| 25 | 2792 | -2.91 | -1.70 | -3.19 | < .001 |
| 26 | 2868 | -2.93 | -1.67 | -3.54 | < .001 |
| 27 | 2942 | -2.95 | -1.74 | -3.80 | < .001 |
| 28 | 3016 | -2.92 | -1.78 | -3.35 | .002 |
| 29 | 3092 | -2.85 | -1.82 | -3.09 | .004 |
| 30 | 3166 | -2.80 | -1.82 | -3.11 | .004 |
| 31 | 3242 | -2.72 | -1.69 | -2.93 | .006 |
| 32 | 3316 | -2.67 | -1.64 | -2.67 | .01 |
| 33 | 3390 | -2.74 | -1.72 | -2.91 | .006 |
| 34 | 3466 | -2.80 | -1.78 | -3.41 | .002 |
| 35 | 3540 | -2.83 | -1.73 | -3.44 | .002 |
| 36 | 3616 | -2.88 | -1.68 | -3.04 | .005 |
| 37 | 3690 | -2.90 | -1.70 | -2.81 | .009 |
| 38 | 3764 | -2.90 | -1.75 | -2.81 | .009 |
| 39 | 3840 | -2.89 | -1.86 | -2.94 | .006 |
| 40 | 3914 | -2.86 | -1.88 | -3.00 | .005 |
| 41 | 3990 | -2.82 | -1.84 | -2.85 | .008 |
| 42 | 4064 | -2.85 | -1.80 | -2.83 | .008 |
| 43 | 4138 | -2.87 | -1.75 | 2.90 | .007 |
| 48 | 4512 | -2.84 | -1.34 | -2.64 | .01 |
| 55 | 5036 | -2.88 | -1.45 | -2.55 | .01 |

| | | | | | |
|---|---|---|---|---|---|
| 56 | 5112 | -2.92 | -1.34 | -2.66 | .01 |
| 57 | 5186 | -2.89 | -1.38 | -2.63 | .01 |

Appendix 5.1. The NASA Task Load Index (NASA-TLX)

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Appendix 5.2. Situation Awareness Rating Technique (SART)

**1. How changeable is the situation? [Instability]**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| stable and straightforward | 1 | 2 | 3 | 4 | 5 | 6 | 7 Changing suddenly |

**2. How many variables are changing within the situation? [Variability]**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Very few variables changing | 1 | 2 | 3 | 4 | 5 | 6 | 7 A large number of factors varying |

**3. How complicated is the situation? [Complexity]**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Simple and straightforward | 1 | 2 | 3 | 4 | 5 | 6 | 7 Complex with many interrelated components |

**4. How aroused are you in the situation? [Arousal]**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| A low degree of alertness | 1 | 2 | 3 | 4 | 5 | 6 | 7 Alert and ready for activity |

**5. How much mental capacity do you have to spare in the situation? [Spare capacity]**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Nothing to spare at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 Sufficient to attend to many variables |

**6. How much are you concentrating on the situation? [Concentration]**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Focusing on only one | 1 | 2 | 3 | 4 | 5 | 6 | 7 Concentrating on many aspects of the situation |

**7. How low much is your attention divided in the situation? [Attention division]**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Focusing on only one | 1 | 2 | 3 | 4 | 5 | 6 | 7 Concentrating on many aspects of the situation |

**8. How much information have you gained about the situation? [Quantity]**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Very little | 1 | 2 | 3 | 4 | 5 | 6 | 7 A great deal of knowledge |

**9.How good information have you been accessible and usable? [Quality]**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Difficult to get required operating parameters / symptoms | 1 | 2 | 3 | 4 | 5 | 6 | 7 Required operating parameters / symptoms are adequately supplied |

**10. How familiar are you with the situation? [Familiarity]**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| New situation | 1 | 2 | 3 | 4 | 5 | 6 | 7 A great deal of relevant experience |

Appendix 5.3. Trust in Automation (TiA) questionnaire

| | | Strongly disagree | Rather disagree | Neither disagree nor agree | Rather agree | Strongly agree | No response |
|---|---|---|---|---|---|---|---|
| 1 | The system is capable of interpreting situations correctly. | ① | ② | ③ | ④ | ⑤ | ○ |
| 2 | The system state was always clear to me. | ① | ② | ③ | ④ | ⑤ | ○ |
| 3 | I already know similar systems. | ① | ② | ③ | ④ | ⑤ | ○ |
| 4 | The developers are trustworthy. | ① | ② | ③ | ④ | ⑤ | ○ |
| 5 | One should be careful with unfamiliar automated systems. | ① | ② | ③ | ④ | ⑤ | ○ |
| 6 | The system works reliably. | ① | ② | ③ | ④ | ⑤ | ○ |
| 7 | The system reacts unpredictably. | ① | ② | ③ | ④ | ⑤ | ○ |
| 8 | The developers take my well-being seriously. | ① | ② | ③ | ④ | ⑤ | ○ |
| 9 | I trust the system. | ① | ② | ③ | ④ | ⑤ | ○ |
| 10 | A system malfunction is likely. | ① | ② | ③ | ④ | ⑤ | ○ |
| 11 | I was able to understand why things happened. | ① | ② | ③ | ④ | ⑤ | ○ |
| 12 | I rather trust a system than I mistrust it. | ① | ② | ③ | ④ | ⑤ | ○ |
| 13 | The system is capable of taking over complicated tasks. | ① | ② | ③ | ④ | ⑤ | ○ |
| 14 | I can rely on the system. | ① | ② | ③ | ④ | ⑤ | ○ |
| 15 | The system might make sporadic errors. | ① | ② | ③ | ④ | ⑤ | ○ |
| 16 | It's difficult to identify what the system will do next. | ① | ② | ③ | ④ | ⑤ | ○ |
| 17 | I have already used similar systems. | ① | ② | ③ | ④ | ⑤ | ○ |
| 18 | Automated systems generally work well. | ① | ② | ③ | ④ | ⑤ | ○ |
| 19 | I am confident about the system's capabilities. | ① | ② | ③ | ④ | ⑤ | ○ |

Appendix 5.4. Movement index. Movement index was calculated to determine whether there was a difference in wrist movement between the expected and unexpected stop. The times of interest were the same as the physiological analyses: 30 seconds before the stop and 30 seconds after the stop. Movement index was calculated by subtracting the x, y, and z coordinates from the previous coordinates and then deriving the Euclidian distance between them. To compare the movement between stops, a 2 (Stop: unexpected, expected) x 2 (Time: pre-stop, post-stop) repeated measures ANOVA was conducted. The results revealed no significant differences between the main effects of stop and time, or the interaction effect ($F_{(1, 29)} \leq 2.10$, $p \geq .159$). Movement was similar before the expected stop ($M$ = 0.03, $SD$ = 0.01), after the expected stop ($M$ = 0.02, $SD$ = 0.02), before the unexpected stop ($M$ = 0.03, $SD$ = 0.01), and after the unexpected stop ($M$ = 0.03, $SD$ = 0.01).

Appendix 6.1. Likert-style questions administered.

| Dimension | Measure | Question |
|---|---|---|
| Usefulness | Usefulness, Satisfaction, and Ease of use (USE) | It helps me be more effective<br>It helps me be more productive<br>It is useful<br>It gives me more control over the activities in my life<br>It makes the things I want to accomplish easier to get done<br>It saves me time when I use it<br>It meets my needs<br>It does everything I would expect it to do |
| Anxiety | Unified Theory of Acceptance and Use of Technology (UTAUT) | I feel apprehensive about using the system<br>The system is somewhat intimidating to me |
| - | In-house | I would have liked to control and changed the speed of the vehicle myself<br>I liked that the system took control and changed the speed of the vehicle<br>I would want to decide when the system was turned on and off<br>I would want options if the driving style was going to change<br>I liked that the system alerted me via the dashboard on screen and with a voice notification<br>I would want to know how it works before using the system |

Appendix 6.2. Movement index. Movement index was calculated to determine whether there was a difference in wrist movement between the driver state monitoring and speed limit change notification. The times of interest were the same as the physiological analyses: 30 seconds before the notification and 30 seconds after the notification. Movement index was calculated by deriving the Euclidean distance between the previous x, y, and z coordinates. To compare the movement between notifications, a 2 (Notification: driver state monitoring, speed limit change) x 2 (Time: pre-, post-) repeated measures ANOVA was conducted. The results revealed a significant difference between the main effect of Journey, $F_{(1, 27)} = 19.51$, $p < .001$, $n_p^2 = .415$. The main effect for Time and the Interaction effect were not significant ($F_{(1, 27)} \leq 0.15$, $p \geq .357$). Movement was greater during the speed limit change notification ($M = 2.24$, $SD = 1.35$) compared to the driver state monitoring notification ($M = 1.14$, $SD = 0.76$).