UNIVERSITY OF GLOUCESTERSHIRE

A SYSTEMATIC REVIEW:

INVESTIGATING THE USE OF CURRENT MEASUREMENT TECHNIQUES FOR THE DEVELOPMENT OF AN OBJECTIVE COGNITIVE FATIGUE METRIC

A. Dickens

MSc by Research

2024

A SYSTEMATIC REVIEW:

INVESTIGATING THE USE OF CURRENT MEASUREMENT TECHNIQUES FOR THE DEVELOPMENT OF AN OBJECTIVE COGNITIVE FATIGUE METRIC

Alicia Dickens

A thesis submitted to the University of Gloucestershire in accordance with the requirements for the degree of Masters by Research in the School of Psychological Sciences

January 2024

Word Count: 32,220

ABSTRACT

Cognitive fatigue is a complex psychobiological state that usually presents as a symptom or as a common occurrence alongside various, highly prevalent conditions and illnesses. As a result, the effects of cognitive fatigue, including a lack of concentration, attention, motivation and lethargy, can greatly impact quality of life and ability to work. Self-report measures, which are important for the understanding of an individual's subjective perspective, are currently used within clinical settings to gauge the impact of cognitive fatigue on a day-to-day basis. However, subjective feelings of cognitive fatigue can, for a variety of reasons, vary from an individual's ability to perform a task well, which stems from their objective level of cognitive fatigue. Therefore, pursuing and establishing a widely used and accessible objective cognitive fatigue measure would be invaluable when considering return-to-work protocols and objective comparisons between participant and/or patient populations. To begin this process, identifying components of a reliable, objective cognitive fatigue measure is paramount. To this end, the aim of this systematic review is to identify and evaluate methods already being used throughout literature for the measurement of objective cognitive fatigue levels. The culmination of appropriate measurement techniques will allow for suggestions for future research, as well as providing comprehensive guidance to other researchers looking to incorporate a measurement technique within their own objective cognitive fatigue studies.

DECLARATIONS

I declare that the work in this thesis was carried out in accordance with the regulations of the University of Gloucestershire and is original except where indicated by specific reference in the text. No part of the thesis has been submitted as part of any other academic award. The thesis has not been presented to any other education institution in the United Kingdom or overseas.

This thesis has been amended for publication. The article will be submitted for publication following the final submission of this document for examination.

Any views expressed in the thesis are those of the author and in no way represent those of the University.

Signed

[redacted]

Alicia Dickens

January 2024

DOI: 10.46289/YYSF3802

ACKNOWLEDGEMENTS

I would first like to express my deepest gratitude and appreciation to my supervisor, Dr Kim Schenke, for being an unwavering source of positivity, productivity and drive. As ever, the completion of this review would not have been possible without your dedication to the practicalities and logistics of the project. Your guidance, patience and support over the past three years throughout my entire journey at the University of Gloucestershire has been second to none, and I cannot thank you enough.

To Dr Andy Champion, who has helped so much with the practicalities of completing the review and always managing to stay upbeat, and to Professor Graham Edgar for his helpful interjections along the way – thank you.

Thank you to my husband, Scott, and our families, for all the love, support, encouragement and comfort that they have provided throughout this process.

Finally, to our very soon-to-be new addition, Oliver, for the motivation, perspective and companionship.

This project was made possible thanks to the funding provided The Research Priority Area in Health, Sports, Wellbeing and Life Sciences at the University of Gloucestershire.

CONTENTS

List of Tables

LIST OF FIGURES

Chapter One: An Introduction to Cognitive Fatigue

Defining & Categorising Cognitive Fatigue

Cognitive Resources & Motivation – The Equilibrium between Objective and Subjective Cognitive Fatigue

The term 'cognitive fatigue' refers to a psychobiological state, which results in an inability to maintain task performance for prolonged periods (Benoit et al., 2017). Cognitive fatigue is hypothesised to be a result of decreased levels of cognitive resources, including a failure to allocate available resources (Proost et al., 2022), known as objective cognitive fatigue. Failure to allocate available cognitive resources occurs mainly through a lack of motivation (Herlambang et al., 2019), or subjective cognitive fatigue. The term 'cognitive resources' refers to the strictly biological components that the brain requires to function in terms of energy, including brain blood flow, oxygen levels, and glucose levels (Watts et al., 2018). These cognitive resources have been shown to decrease for a number of reasons, resulting in low energy, and therefore cognitive fatigue. In neurotypical individuals cognitive fatigue will occur as a result of day-to-day tasks at work and at home (Walker et al., 2019).

Cognitive fatigue is suggested to be a correlate of inability to maintain one's attention (Chaudhuri & Behan, 2004). Both motivation decline and resource depletion impacts attentional skills, and vice versa, with prolonged attention reducing motivation and cognitive resources (Reteig et al., 2019). As a result of the decrease in motivation and resources, attentional skills and vigilance may also decline, which is where issues surrounding competent decision making and safety arise. There are two main theories for attention resource allocation leading to a disruption in vigilance. The first is the Resource Theory (Davies & Parasuraman, 1982), which states that the need for sustained attention during a prolonged task requires a high level of cognitive effort to stay continually engaged. This consequently depletes information processing resources and leads to a disruption in sustained attention (Davies & Parasuraman, 1982). This loss in attention over time is referred to as a vigilance decrement (Thiffault & Bergeron, 2003). The second theory is the Mindlessness Theory (Manly, 1999), which contrastingly hypothesises that attentional loss is not because of resource depletion. It is, instead, due to the mundane nature of the task, and therefore, a lack of motivation to commit to an appropriate response. This is because of an increased proportion of non-target stimuli that results in a higher level of inhibition needed in order to respond only to the less frequent target stimuli appropriately (Manly, 1999).

The more recent Resource-Control Theory aims to combine both the Resource and Mindlessness Theory. It has been noted that neither Resource nor Mindlessness Theory alone can adequately justify the variety of empirical findings (Thomson et al., 2015). The Resource-Control Theory postulates that effortful executive function to maintain task attention, i.e., cognitive resources, are required and then become depleted over time. The depletion of executive function ("control") consequently results in attentional loss as it becomes too costly to continue allocating resources to the given task. Therefore, motivation decreases as effort increases, and cognitive resources are consequently shifted to off-task, less effortful distractions.

Motivation is a neuroeconomic process of effort distribution allowing for reallocation of cognitive resources based on the individual's subjective opinion of the importance of a task (Herlambang et al., 2020). Limited resources result in the individual being less likely to begin or complete the task (Westbrook & Braver, 2015). Effort distribution can be shifted, with cognitive resources reallocated when an individual deems a task to be worth achieving (Behrens et al., 2023). The Motivational Control Model (Hockey, 2011) suggests that cognitive resource allocation can shift depending on the demands of a task, with goal management as the predominant factor in resource allocation. Goal importance is determined subjectively by intrinsic or extrinsic motivational factors (Hockey, 2011). Extrinsic motivational factors include tangible reward, such as a monetary incentive. Comparatively, intrinsic motivation comes from innate sensations of achievement, competition, purpose or recognition (Baldassarre, 2011). This model is therefore dependent on the individual's subjective belief that the goal is worth the investment of cognitive resources, and therefore demonstrates that cognitive fatigue can be impacted on a subjective level.

The combination of motivation and cognitive resources is why it is appropriate to refer to cognitive fatigue as a psychobiological state. I propose that the differing levels of cognitive resources and subjective motivation give rise to four overarching outcomes, as demonstrated below in Figure 1.

Figure 1

A punnet square diagram depicting the relationship between motivation and cognitive resources

Availability of cognitive resources + motivation **= no cognitive fatigue**	Availability of cognitive resources + no motivation **= boredom/task aversion/subjective cognitive fatigue**
Decreased availability of cognitive resources + motivation **= objective cognitive fatigue**	Decreased availability of cognitive resources + no motivation **= burnout/apathetic/attentional loss through extreme objective and subjective cognitive fatigue**

Recent studies have revealed a disparity between subjective and objective cognitive fatigue when the effort-reward balance of a task becomes unfavourable (de Jong et al., 2018). This causes an increase in subjective cognitive fatigue despite adequate objective task performance. This discrepancy is attributed to individuals evaluating the trade-off between effort and reward subjectively, leading to heightened perceptions of fatigue, while objective task performance remains satisfactory (Ramage et al., 2019; Ricci et al., 2007). For example, individuals are cognitively able to complete a task, but take longer or find it more difficult to do so due to lack of motivation. Conversely, a misalignment between subjective and objective cognitive fatigue can also arise in situations of heightened motivation when coupled with diminished cognitive resources. This results in subjective cognitive fatigue levels indicating less cognitive fatigue than objective measures suggest (Ramage et al., 2019; Ricci et al., 2007). In this instance, even though individuals may successfully complete a task, it may take more time or prove more challenging than it would normally be given adequate cognitive resources, despite high motivation. The prolonged time-on-task, coupled with increased mental effort, accelerates the depletion of already depleted cognitive resources, leading to the need for extended recovery periods.

Thus, cognitive fatigue can consist of both the subjective experience of cognitive fatigue and the actual objective decline in task performance. With the outcomes from the misalignments between subjective and objective cognitive fatigue presenting as extremely similar regardless of the dominant factor, it is unsurprising that existing literature often discusses cognitive fatigue as a singular state. This results in research neglecting distinct mechanisms and measurement techniques for the separate assessment of subjective and objective cognitive fatigue (Manjaly et al., 2019). Complicating matters further, literature frequently encompasses related concepts like attention, concentration, drowsiness, and boredom under the overarching term of cognitive fatigue, leading to a lack of clarity in research. These entities are closely intertwined with cognitive fatigue, to the extent thar terms such as "cognitive workload" and "cognitive effort" are used interchangeably, and outcome measures heavily rely on participants' attention, drowsiness, concentration or alertness to determine overall cognitive fatigue levels. Although these entities are relevant to cognitive fatigue measurement, appropriate subjective measures should be provided alongside any objective testing. This approach would help identify the influence of subjective factors such as motivation, sleep deprivation, boredom, and substance intake (e.g., caffeine and sugar) on overall cognitive fatigue levels. Consequently, integrating subjective measures will contribute to standardising objective cognitive fatigue assessments in research, both when conducting research and compiling data.

Pathological Cognitive Fatigue and Impact on Life

Both general and pathological cognitive fatigue exhibit a diverse range of symptoms, encompassing but not limited to diminished motivation, reduced attention and concentration, heightened physical lethargy, and increased exhaustion (Mendoza-Ruvalcaba et al., 2019; Slimani et al., 2018). When it comes to pathological cognitive fatigue, whether as a standalone condition or comorbidity, it can continue to affect individuals chronically ('chronically' being defined as a minimum of six months), significantly impacting individuals' ability to carry out routine tasks and diminishing their overall quality of life (Palacios et al., 2023; Walker et al., 2019). When cognitive fatigue has a pathological aetiology, the initial levels of cognitive resources needed to complete everyday tasks are already diminished (Rönnbäck & Johansson, 2022).

Approximately 20-24% of the population express feelings of cognitive fatigue, often occurring alongside or as a symptom of other conditions such as stroke, multiple sclerosis and acute brain injury (Dantzer et al., 2014). Consequently, it has become one of the most prevalent reasons for seeking advice from healthcare professionals (Harrington, 2012; Sabes-Figuera et al., 2010; Yoon et al., 2023). In pathological terms, it stands to reason that biological resources are likely to be redirected to help the body heal, resulting in cognitive fatigue being a prevalent comorbidity. Rink & Khanna (2011) determined that, during periods of rest, the brain requires 20% of total body oxygen for processes such as ATP production through oxidation of glucose. The proportion of overall body weight, of which the brain accounts for approximately 2%, in comparison to oxygen consumption is disproportionately high. The long-term effects of reduced oxygen levels are notably recorded in stroke patients, with 30-70% of survivors reporting cognitive fatigue, regardless of whether permanent neurological impairments are present or not (Johansson & Rönnbäck, 2012). The human brain also requires 20% of the total body glucose consumption for processes such as ATP production, neurotransmitter and neuromodulator synthesis, and oxidative stress control. Thus, a deficit in glucose can result in impaired brain function, manifested as cognitive fatigue (Mergenthaler et al., 2013; Mink et al., 1981). This is evident in cases of mild traumatic brain injuries, where 68% of patients report cognitive fatigue one week after the injury, with 34% continuing to experience it six months later (Norrie et al., 2010).

While acute cognitive fatigue in healthy individuals typically quickly dissipates after rest (Mizuno et al., 2011), disparities between subjective and objective cognitive fatigue become more significant in cases of sudden changes in fatigue states, such as traumatic brain injury or stroke. In such instances, individuals may be unaware or in the process of understanding their own experience of cognitive fatigue (Ramage et al., 2019). Consequently, the presence of cognitive fatigue as a secondary symptom of disease or trauma is concerning, as it often entails an unnoticed misalignment between objective and subjective cognitive fatigue levels, thereby inadvertently increasing the risk of accidents or injuries (Ramage et al., 2019; Ricci et al., 2007). Research into the separate categorisation of cognitive fatigue has improved in recent years to find solutions to the positive correlation between increased objective cognitive fatigue and risk of accident and injury (Proost et al., 2022). The ability to accurately assess cognitive fatigue levels objectively is being more recognised as crucial for workplace safety and may help individuals return-to-work in a timelier manner (Guo et al., 2016).

The impact of cognitive fatigue extends to driving capability and employment status: cognitive fatigue is considered to be a predominant cause of work-related and road traffic accidents (Turner, 2013), and several studies demonstrate a relationship between prolonged cognitive fatigue and unemployment (Hiploylee et al., 2017; Losoi et al., 2016). Importantly, it is not just the time at work that needs to be considered, but also the journeys to and from work. For example, for some individuals, driving to work may be a considerably more dangerous task than for others, particularly during longer journeys where sustained attention is necessary (Ma et al., 2018; Massar et al., 2018). Indeed, road traffic accidents are one of the main non-health related causes of death, with over 1.35 million people killed in traffic accidents each year around the world (Ahmed et al., 2023; Ma et al., 2018). Hypo-vigilance and fatigue have been regarded as significant contributors for road accidents. For example, empirical studies have reported that approximately 10–20% of car crashes on the highways in the southwest and the midlands in England were caused by sleepiness or fatigue (Moradi et al., 2019). Albeit precise contributors of these factors have not yet been determined in crashes, researchers have reached a consensus that fatigue represents a major road safety hazard (Moradi et al., 2019). Furthermore, the social and economic cost of traffic incidents has been estimated to total a global cost of US$518 billion per year (Jimenez-Pinto & Torres-Torriti, 2013). Thus, developing accurate measures and coping strategies for managing cognitive fatigue is crucial within driving to reduce road traffic collisions (Ma et al., 2018).

The inability to return-to-work not only impacts quality of life on an individual level, but also has a larger scale financial impact in the workforce sector. Lost productive hours as a result of workers with performance-related fatigue (a decline in performance due to physical or cognitive fatigue), as well as hidden informal care costs, are estimated at £3.5 billion for the UK (Sabes-Figuera et al., 2010). Indeed, individuals returning to work after sick leave for mental health reasons (i.e., due to cognitive fatigue or mental illness) are more likely to work reduced hours or leave employment in comparison to workers who have not previously taken mental health sick leave (Norder et al., 2016). Furthermore, motivational levels can impact cognitive fatigue. This can consequently result in decreased attention, alertness, and mental efficiency (Muñoz-de-Escalona et al., 2020). For this reason, subjective cognitive fatigue can result in individuals leaving work, and feeling unable to return-to-work either at all or in a limited capacity, whilst objective cognitive fatigue can result in serious errors and diminished job performance (Dawson et al., 2011; Hopstaken, 2016). For those who have been subject to drastic changes in cognitive resources due to illness or injury, becoming accustomed to the debilitating effects of objective cognitive fatigue and the limitations due to reduced cognitive resources can present as a hugely personal and frustrating challenge (Tomar et al., 2018). Furthermore, symptoms of cognitive fatigue overlap with symptoms of mental illnesses, oftentimes presenting as depression, which can lead to wrongly diagnosed mental health presentations (Billones et al., 2020; Tomar et al., 2018). Two-thirds of individuals who suffer with depression are granted sick leave (Norder et al., 2016), with mental health illness, and depression in particular, accounting for the majority of sick leave and unemployment (Porter et al., 2018). However, if the correct diagnosis is in fact cognitive fatigue, different interventions would be required. Additionally, being unable to work is associated with lower self-esteem, due to the inability to contribute financially, socially or societally, which in turn can worsen mental health symptoms, including cognitive fatigue (Álvaro et al., 2019). This significant personal and economic impact highlights the need to assess current return-to-work protocols and the support provided as affected individuals transition back to the workplace.

An Example: Stroke

In the UK, return-to-work is achieved by less than 50% of stroke survivors (Radford et al., 2020). Furthermore, stroke prevalence is rising in younger people, with individuals under 69 years accounting for approximately 41% of stroke survivors (Public Health England, 2018). It is estimated that two-thirds of stroke survivors have a residual physical or cognitive disability that limits their ability to return-to-work, and further factors such as employer attitudes and timely access to rehabilitation and support programs can also result in decreased chances of a successful reintegration (Coole et al., 2013). The economic burden of stroke for health and social services is estimated at £26 billion annually, with informal and ongoing care thought to cost a further £20.6 billion (Patel et al., 2020). Returning to work is one way in which this economic burden can be diminished. Work provides a source of social interaction, identity and purpose as well as supporting health, wellbeing and longevity. This in turn decreases lost productivity hours, reduces the risk of depression, improves quality of life and reduces other unhealthy behaviours that may otherwise add to ongoing care costs (Thill et al., 2020). However, stroke survivors are often discharged from hospital without further support or guidance regarding return-to-work, with more severe cases prompting healthcare professionals to prematurely recommend not returning to work based on the disabilities and the individual's capacity to fulfil the requirements of their previous occupation (Green et al., 2021). Furthermore, return-to-work services are not widely available, with only 37% of primary care trusts providing vocational rehabilitation (Sinclair et al., 2014). Where these services are accessible, they are disorganised due to the lack of a sanctioned pathway for support, with mild stroke patients receiving little to no support (Radford et al., 2020; Sinclair et al., 2014).

The individual impact of a stroke on any given stroke survivor is incredibly varied, and cognitive fatigue is one of many difficulties a stroke survivor may face on the road to recovery. If a stroke survivor has obtained more severe disabilities, cognitive fatigue is unlikely to be seen as a priority (Saar et al., 2023; Sinclair et al., 2014). In milder cases, cognitive fatigue is still highly reported and has the potential to be addressed as necessary. This is crucial, as cognitive performance has been found to be significantly correlated with returning to work (Saar et al., 2023). However, even in more recent studies assessing cognitive performance for return-to-work, the focus of the assessments is based predominantly in subjective measures, with questionnaires assessing self-perceived working ability, motivation and barriers, mood, self-perception, memory and attention (Barboza et al., 2017; Saar et al., 2023; Tay et al., 2023). Although an important consideration, this method of assessment relies solely on self-perceived task performance, rather than on objective measure of task performance that can be correlated to objective cognitive fatigue. Based on the above theories and evidence, these protocols are therefore missing a crucial aspect of being able to determine an individuals' ability to return-to-work successfully.

It is therefore essential to accurately evaluate both objective and subjective levels of cognitive fatigue to facilitate individuals' timely return-to-work without exceeding their current capabilities. It is crucial that these measures enable and empower individuals to actively monitor their own progress as they reintegrate into work, allowing them to take breaks when necessary and appropriate (Guo et al., 2016). This becomes especially important in occupations with heightened risk, whereby cognitive fatigue-induced attentional impairments could pose potentially life-threatening consequences. Examples of such occupations include operating machinery, driving long distances and making critical decisions in healthcare settings (Dawson & Fletcher, 2001; Guo et al., 2016). Professions such as power plant workers, air traffic controllers, pilots, long-distance drivers and military personnel also demand a high level of sustained attention for extended periods in relatively high-risk situations (Gartenberg et al., 2018). A measure that could allow for employers and healthcare professionals to track objective cognitive fatigue levels in a collaborative manner would add to creating a well-supported and team-oriented approach to the individuals' wellbeing and safety at work.

CURRENT MEASUREMENT TECHNIQUES

Currently, the majority of measures focus on subjective cognitive fatigue through self-report measures. For example, questionnaires are employed in order to determine the individual's subjective level of cognitive fatigue, including Likert-style self-report measures (Ashman et al., 2008), the Patient-Reported Outcomes Measurement Information System (PROMIS), the Quality of Life in Neurological Disorders (Neuro-QOL) system, the Visual Analogue Scale, the NASA-Task Load Index and the more newly developed Traumatic Brain Injury Quality of Life Fatigue item bank (Kisala et al., 2019).

Despite the utilisation of a variety of questionnaires, it is important to find a way to assess cognitive fatigue objectively due to the psychobiological nature of cognitive fatigue discussed earlier: motivation and cognitive resources play a huge role in cognitive fatigue levels and measuring only subjective cognitive fatigue does not capture the full picture. As a result of this, more recent literature has been exploring ways to objectively measure cognitive fatigue. There are two key approaches to measure cognitive fatigue objectively; physiological or task performance based. Physiological measures include neuroimaging such as electroencephalogram (EEG), functional magnetic resonance imaging (fMRI) and blood oxygen level dependent (BOLD) imaging, electrocardiogram (ECG), eye-tracking and pupillometry, and galvanic skin conductance (Cook et al., 2007; Zhang et al., 2021). In comparison, task performance measures utilise response times, response rate, response accuracy and time-on-task. There are a number of advantages and disadvantages of both physiological and task performance measures, which will also be discussed here briefly following an overview of how these measures can be utilised.

Event-related potentials are small electrical signals generated in specific brain structures as a result of specific stimuli, and can be captured during EEG as a correlate of objective measurement of task performance (Petruo et al., 2018). Event-related potentials are useful for cognitive fatigue measurement because the positive (Px) and negative (Nx) waves that are associated with different areas of brain function can be tracked and analysed with regards to the task that is being undertaken (Möckel et al., 2015). Changes in amplitude and latency of the waves have been studied in order to determine any relation to cognitive fatigue. However, there is a lack of consensus on exactly which components are the most optimal indicators of cognitive fatigue. For example, Xiao et al. (2019) found that N1, P2 and P3 waves, which are thought to be associated with working memory and attention (Assecondi et al., 2022), decrease in amplitude during cognitive fatigue, whereas Möckel and colleagues (2015) did not find an N1 amplitude decrease. Meanwhile, N2 and P2 latency were shown to be increased in a cognitively fatigued state following prolonged time-on-task (Möckel et al., 2015; Xiao et al., 2019). Therefore, further research is needed to fully understand the interplay between the various underlying components and cognitive fatigue.

Blood oxygen level dependent (BOLD) activation can also be used to objectively identify areas of the brain that are functioning during tasks and whether the level of functioning changes over a period of time (Genova et al., 2013). Slow BOLD activation in response to tasks has been reported alongside recruitment of wider brain regions (Mizuno et al., 2015; Shan et al., 2018), as well as an increase in BOLD response with no significant increase in performance levels in comparison to control groups (Lange et al., 2005; Mizuno et al., 2015; van der Schaaf et al., 2018). However, the data collected from BOLD activation is unlikely to be specific enough for same-day changes in objective cognitive fatigue levels (Iancheva et al., 2019). Similarly, data collected from fMRI does not provide a sensitive enough measure: although brain activity, and therefore, where the cognitive resources are required, can be discovered for different tasks, it is not a sensitive enough measure to discriminate between different activity levels throughout the day (Cook et al., 2007). Furthermore, the technology needed to access this data would not be viable in terms of accessibility and expense for daily measurement of cognitive fatigue for use in the workplace.

Heart rate variability has also been shown correlate with cognitive fatigue, making ECG a viable objective measure. This stems from the autonomic nervous system having a key part in behaviour adaptation and modulation during long, high cognitive effort tasks (Giuliano et al., 2018; Lin et al., 2020; Matuz et al., 2021). Heart rate variability is an indicator of cardiac autonomic regulation that is linked to a variety of brain regions, including the striatal and limbic regions, which play a role in parasympathetic autonomic modulation (Matuz et al., 2021; Mizuno et al., 2014). Parasympathetic modulation has been shown to decrease with sympathetic modulation and increase when challenging cognitive tasks are being undertaken. This correlate is well established with cognitive fatigue within the literature (Gergelyfi et al., 2015; Matuz et al., 2021; Mizuno et al., 2011, 2014; Segerstrom & Nes, 2007; Tran et al., 2009). Heart rate variability is therefore thought to be a reliable measure for determining the link between cognitive load and autonomic modulation (Matuz et al., 2021). Zhong and colleagues (2005) also found that decreased parasympathetic and increased sympathetic modulation was present within sleep deprived individuals. This study showed a much more consistent significant difference in respiratory rate, with no significant difference between heart rate for all three cognitive tests undertaken (2-back task, shape Go/NoGo and simple reaction time task). This is important to note as it decreases the likelihood that heart rate is measuring a potential confounding factor of sleepiness, which could be advantageous or disadvantageous, depending on the application. However, heart rate can increase or decrease for a number of other reasons than just cognitive fatigue. For example, if the person is excited, afraid or stressed (Abdullah & Mustapar, 2019). This means caution must be taken when using heart rate as a measure to ensure it is capturing cognitive fatigue and not something else.

Galvanic skin conductance has been shown to distinguish between rest and task conditions, however, there has been conflicting evidence as to its sensitivity. Widyanti and colleagues' (2017) study determined that galvanic skin conductance is not sensitive enough to distinguish between levels of cognitive fatigue, but previous literature has found a correlation between galvanic skin conductance and different levels of cognitive fatigue (Engström et al., 2005; Ikehara & Crosby, 2005). Nourbakhsh and colleagues (2012) investigated galvanic skin conductance using a variety of arithmetic and reading levels and found that applying normalisation to their results cancelled out subject-dependency within the data. This revealed a significant distinction between the cognitive load induced by different task difficulty levels and the mean and accumulative galvanic skin conductance. Therefore, galvanic skin conductance could be a non-invasive, physiological measure for cognitive load, which has been shown to reflect levels of cognitive fatigue (Mizuno et al., 2011; Tanaka et al., 2014). The correlation between these factors would then need to be discussed to assure research transparency, as other confounding factors can play a part in causing physiological responses.

Eye-tracking and pupillometry is also a common objective cognitive fatigue method, as it is relatively easy to set-up due to advances in technology and has a particular appeal for driving fatigue experiments. However, comparisons between studies have been made unclear through the use of varied analysis techniques, equipment, experimental methods and lighting constraints (Bafna & Hansen, 2021). This means that finding specific, reliable correlates between eye-tracking and pupillometry data that would also be able to reach a high sensitivity level is difficult. Additionally, although the use of eye-tracking could be implemented relatively easily in an office setting, it would not be particularly accessible for individuals who do not have consistent access to a computer as part of their occupation.

Tate and colleagues (2013) found in their systematic review that the most commonly employed performance-based executive test for assessing traumatic brain injury in adults was the

Wisconsin Card Sorting Test (WCST), and the most common executive skills battery was the Behavioural Assessment of the Dysexecutive Syndrome (BADS). Executive performance is affected by cognitive resource allocation, and therefore, cognitive fatigue levels. Similarly, tasks such as the n-back, time-load dual-back task and the Stroop task have also been used to assess task performance by measuring the change in fatigue level over time, or as a pre-/post-test assessment of cognitive fatigue (Borragán et al., 2017). However, throughout the literature, studies use these tasks to induce cognitive fatigue, as these tasks increase cognitive load and therefore increase demands on cognitive resources (O'Keeffe et al., 2020). Whilst this is appropriate when simply wanting to objectively measure fatigue changes over the course of a short examination in a hospital or laboratory setting, using a task that is, itself, cognitively fatiguing would not be appropriate for assessing how one's cognitive fatigue fluctuates throughout the day. Yet, as highlighted above, the disconnect between subjectively feeling cognitively fatigued and performance and attention being objectively reduced on a task, to inform return-to-work protocols to maintain workplace safety requires an objective measure that does not in itself contribute to cognitive fatigue. There are some performance-based tasks that sustain less strain on cognitive resources as they do not require use of working memory such as the n-back or time-load dual-back task, and mostly focus on the individuals' attentional skills (Truszczynski et al., 2009; Vrijkotte et al., 2018).. Indeed, as discussed previously, attentional loss occurs when both motivation and cognitive resources are decreased, meaning that in combination with a subjective, motivation-specific questionnaire, use of attention as a primary measure could provide a good indication of objective cognitive fatigue.

RATIONALE AND OVERVIEW

Currently, longitudinal studies, questionnaires and clinical reviews allow for a participant-led response to cognitive fatigue, which is important for exploring the subjective effects of cognitive fatigue. However, objective cognitive fatigue does not currently have a measure that is widely used and available within clinical settings or for individuals to use on a day-to-day basis. Cognitive fatigue levels fluctuate throughout the day due to task demands, such as within a working environment (Hopstaken et al., 2015). Therefore, being able to efficiently assess daily fatigue over a chosen timeframe is vital in order to understand the impact of cognitive fatigue on individuals and to track temporal characteristics that could help to determine any underlying triggers or aetiologies (Christodoulou et al., 2014). An objective measure would also be paramount for accurate estimation of the impact of cognitive fatigue within and between participant/patient groups for monitoring treatment plans, as well as to gain insight into patterns of idiopathic improvement of cognitive fatigue. However, no current measure allows for accurate ‘equivalent’ or objective comparison between participant/patient groups (Theadom et al., 2016). With these considerations in mind, appropriate design concepts for an objective cognitive fatigue measure would need to include making it suitable for use in a variety of workplaces. The measure would also need to be reliable, affordable, and time efficient, so that it could be used on a regular basis. Finally, it would not be appropriate to have a measure that exacerbates cognitive fatigue itself. This would allow for a decline in objective fatigue to be tracked accurately and acted upon accordingly.

This thesis systematically reviews previous research surrounding existing objective measures of cognitive fatigue to identify potential measures that may be appropriate for the daily measurement of cognitive fatigue that could be used to support return-to-work protocols in clinical patients, and which could also be beneficial to the wider workforce. Chapter Two will present the findings of our systematic review with a comprehensive narrative synthesis on existing methods used to induce and measure objective cognitive fatigue. In Chapter Three, the results from the systematic review discussed, taking into consideration the strengths, limitations and next steps of objective cognitive fatigue measurement and research according to these findings. The planned meta-analysis could not be undertaken due to a large proportion of studies not providing effect sizes for their results, highlighting the lack of open science practices in reporting effect sizes alongside inferential statistics.

Chapter Two: Systematic Review

Introduction

The term cognitive fatigue refers to a psychobiological state which hinders the ability to maintain task performance over extended periods (Borragán et al., 2017). This is primarily attributed to its broad array of symptoms, which includes, but is not limited to, a lack of motivation, attention and concentration, heightened physical lethargy, and exhaustion (Mendoza-Ruvalcaba et al., 2019; Slimani et al., 2018). The diverse symptoms associated with cognitive fatigue can significantly impact individuals' quality of life, whether as a coexisting or standalone condition (Walker et al., 2019). Additionally, the multifaceted nature of its symptoms makes investigating the underlying mechanisms behind cognitive fatigue challenging, despite its high prevalence (Dantzer et al., 2014). Moreover, there are currently no evidence-based treatments available for managing the symptoms of cognitive fatigue (Wylie & Flashman, 2017). Existing literature strongly indicates that cognitive fatigue can substantially affect an individuals' quality of life by making it difficult to complete daily tasks (Walker et al., 2019). Furthermore, severe cognitive fatigue can persist for prolonged periods, with numerous studies demonstrating a relationship between chronic cognitive fatigue and difficulties in maintaining or returning to work (Hiploylee et al., 2017; Losoi et al., 2016).

As previously discussed, two main factors are hypothesised to contribute to cognitive fatigue: reduced levels of mental resources, including a failure to appropriately allocate available resources (Proost et al., 2022), and a lack of motivation (Herlambang et al., 2020). Because cognitive fatigue stems from both biological and psychological changes, it can be categorised into 'objective' (biological) or 'subjective' (psychological) states (Benoit et al., 2017). However, research often treats cognitive fatigue as a single state entity, resulting in a lack of exploration into the separate mechanisms, measurement, and assessment of subjective and objective cognitive fatigue is lacking (Manjaly et al., 2019).

Current research indicates that subjective and objective cognitive fatigue arise when the advantages of completing a task are outweighed by the cognitive resources needed (de Jong et al., 2018). High levels of cognitive fatigue detrimentally affects attention, alertness, and concentration levels (Muñoz-de-Escalona et al., 2020), leading to increased interest in this field due to its link with higher risk of accidents and injuries (Proost et al., 2022). The ability to accurately assess cognitive fatigue levels objectively is crucial for workplace safety and may help individuals return-to-work in a timely manner (Guo et al., 2016).

Presently, healthcare providers typically use subjective, Likert type self-report questionnaires to measure cognitive fatigue (Benoit et al., 2017; NICE, 2021). However, there is growing evidence to suggest that this approach may not always provide an accurate assessment. Subjective and objective levels of cognitive fatigue do not always display a positive correlation as one might expect (Ramage et al., 2019). Often, subjective measures can be indicative of increased fatigue, while objective tasks are still completed with an adequate performance level (Ricci et al., 2007). Conversely, subjective measures might indicate less fatigue than objective measures. If this is a discrepancy that is genuinely manifested within individuals, dangerous consequences may occur through poor decision making, such as when driving (Ricci et al., 2007).

While acute cognitive fatigue in healthy individuals tends to quickly resolve with rest (Mizuno et al., 2011), disparities between subjective and objective cognitive fatigue can become significant following sudden changes in fatigue state, as seen in traumatic brain injuries or stroke (Ramage et al., 2019). The individual may be unaware of or in the process of learning about their personal experience of cognitive fatigue as a result of this sudden change in circumstance. Therefore, it is concerning when cognitive fatigue occurs as a secondary symptom of disease or trauma, as there is often a disparity between objective and subjective cognitive fatigue levels that individuals are unaware of, which unknowingly exacerbates the risk of accident or injury (Ramage et al., 2019; Ricci et al., 2007).

The complexity of cognitive functions affected by cognitive fatigue complicates matters, and interchangeable terms for cognitive fatigue such as “cognitive workload” and “cognitive effort” further reduce research clarity. Outcome measures relying on participants’ attention, drowsiness, concentration or alertness also contribute to the challenge of standardising cognitive fatigue research, particularly when “task performance” is considered the definition of objective cognitive fatigue, and there is no accessible measure that is widely available for use by the general public (Linnhoff et al., 2019).

Research Objectives

The objective of this study is to conduct a thorough assessment through critically evaluating existing methods for objectively measuring cognitive fatigue. The aim is to provide insights that can inform the development of a regular and practical objective measure. This would facilitate daily monitoring of individuals' objective cognitive fatigue levels. By employing a systematic review approach in accordance with Prospero and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, we will analyse current techniques for measuring objective cognitive fatigue. The focus will be on evaluating their suitability to be adapted for use in everyday situations, particularly with a focus on improving return-to-work protocols.

METHODS

This review is registered on the PROSPERO database (Dickens et al., 2022). This systematic review was conducted according to Preferred Reporting Items for Systematic review and Meta-Analyses (PRISMA) guidelines and the Cochrane Handbook (Page et al., 2021).

Research Validation

This systematic review compiles studies employing objective measures of cognitive fatigue to help inform the development of a readily available assessment tool for objectively gauging cognitive fatigue, using task performance as an indicator. A particular beneficial characteristic of this task would be the standardisation of task performance measures which can cover a diverse range of applications. A successfully accessible measure could find applications in various settings, giving the prominence of cognitive fatigue in return-to-work protocols, workplace safety, driving, machinery operation and healthcare (Wylie & Flashman, 2017). Additionally, the importance of preventing over-exertion is emphasised, as it is crucial in enabling recovery from chronic cognitive fatigue conditions, injury, and relapses in long-term comorbid conditions (Rudroff et al., 2020). This approach also aims to mitigate the negative impact of mental health and stress-related issues on individuals' personal and professional lives (Dawson et al., 2011; Dawson & Fletcher, 2001). Both of these aspects work symbiotically to increase quality of life: by creating an individualised and collaborative protocol, individuals are more likely to return-to-work in a non-premature, timely manner, which in turn increases chances of successful long-term reintegration into the workforce sector (Saar et al., 2023).

Eligibility Criteria

To determine the keywords for this systematic review, the PICOTS framework (Hartmann et al., 2012), which specifies, *a priori*, criteria based on Population, Intervention, Comparison, Outcome, Time, and Study Design (see *Table 1*). Studies were not excluded based on population (provided they were over 18 and not animal studies), as we had specified the general population as our criteria, and this provided the most thorough search for methods. Our intervention was focused on objective tests of cognitive fatigue, that did not simultaneously induce cognitive fatigue, and were not case studies or pilot studies, as this would allow for more robust sample data. The comparison was made between baseline objective cognitive fatigue levels and post-test objective cognitive fatigue levels. In terms of outcomes, a primary and secondary outcome were defined. The primary outcome was the difference (either in scores/means or effect size) in the objective measure(s) of cognitive fatigue from the baseline. The secondary outcome was the comparison with a subjective fatigue measure. Time was set as a minimum of 30 minutes on task from the initial baseline measure as previous literature has shown a minimum of 30 minutes time-on-task is required to reliably induce cognitive fatigue (Borragán et al., 2017; Holgado et al., 2020; Tanaka, 2015). Any setting was deemed appropriate to ensure a wide and inclusive search. An upload of the systematic review plan can be found pre-published on PROSPERO (Dickens et al., 2022).

Table 1

A table outlining the PICOTS framework

PICOTS Framework	Criteria
P - Population	Humans over 18 years of age
I - Intervention	Induction of objective cognitive fatigue
C - Comparison	Baseline objective cognitive fatigue levels and post-test objective cognitive fatigue levels
O - Outcome	Primary outcome: difference (either in scores/means or effect size) in the objective measure(s) of cognitive fatigue from the baseline, measured with a task that did not simultaneously induce cognitive fatigue Secondary outcome: comparison with a subjective fatigue measure.
T- Time	Minimum of 30 minutes completing a cognitive fatigue inducing task
S – Study Design/Setting	Any

Information Sources and Search Strategy

The subject librarian at the University of Gloucestershire was consulted in the development of the search strategy. Eight electronic databases were searched until December 31, 2022: Scopus; PubMed; PsychArticle; ProQuest; Google Scholar; The Cochrane Library; The Health Technology Assessment Database; and Web of Science. ProQuest was also searched for grey literature in order to maximise results. Table 2 outlines the number of results sourced from each database. Medical subject headings (MeSH) terms were used for the initial literature search. The keywords used for all database searches were: "Mental fatigue" OR "cognitive fatigue" OR "mental exertion" OR "cognitive exertion" OR "cognitive load" AND "objective" AND "measure". A further search of articles on Google Scholar was conducted until December 31, 2023 and yielded no relevant results.

Table 2

A table outlining the number of articles sourced from each database during the search process

Database	Number of Articles Sourced
Scopus	83
PubMed	116
PsychArticle	121
ProQuest	6792
Google Scholar	11,109
The Cochrane Library	652
The Health Technology Assessment	28
Web of Science	4092
ProQuest Grey Literature	552

Study Selection

After completing the initial literature search, all the resulting studies were compiled in Microsoft Excel for duplicate removal. Following this, two reviewers (AD and KS) independently screened the studies on title and abstract against the following hierarchical inclusion criteria - the study must:

1. focus on objective measures of cognitive fatigue in humans;
2. provide original primary data and was not a review;
3. be available in English.

Given the limited information typically provided by titles and abstracts, more detailed criteria (based on the aforementioned eligibility criteria) were not considered appropriate for the initial screening process.

Any conflicts that arose were resolved by a third reviewer (AC). Full texts of the remaining studies were then found and screened by the same two reviewers (AD and KS) against the complete hierarchical exclusion criteria:

1. No access to the paper;
2. No primary data for objective cognitive fatigue;
3. A case study, pilot study or proof of concept;
4. Not a separate task for inducing and measuring cognitive fatigue;
5. The task used to induce cognitive fatigue did not reach a 30-minute threshold.

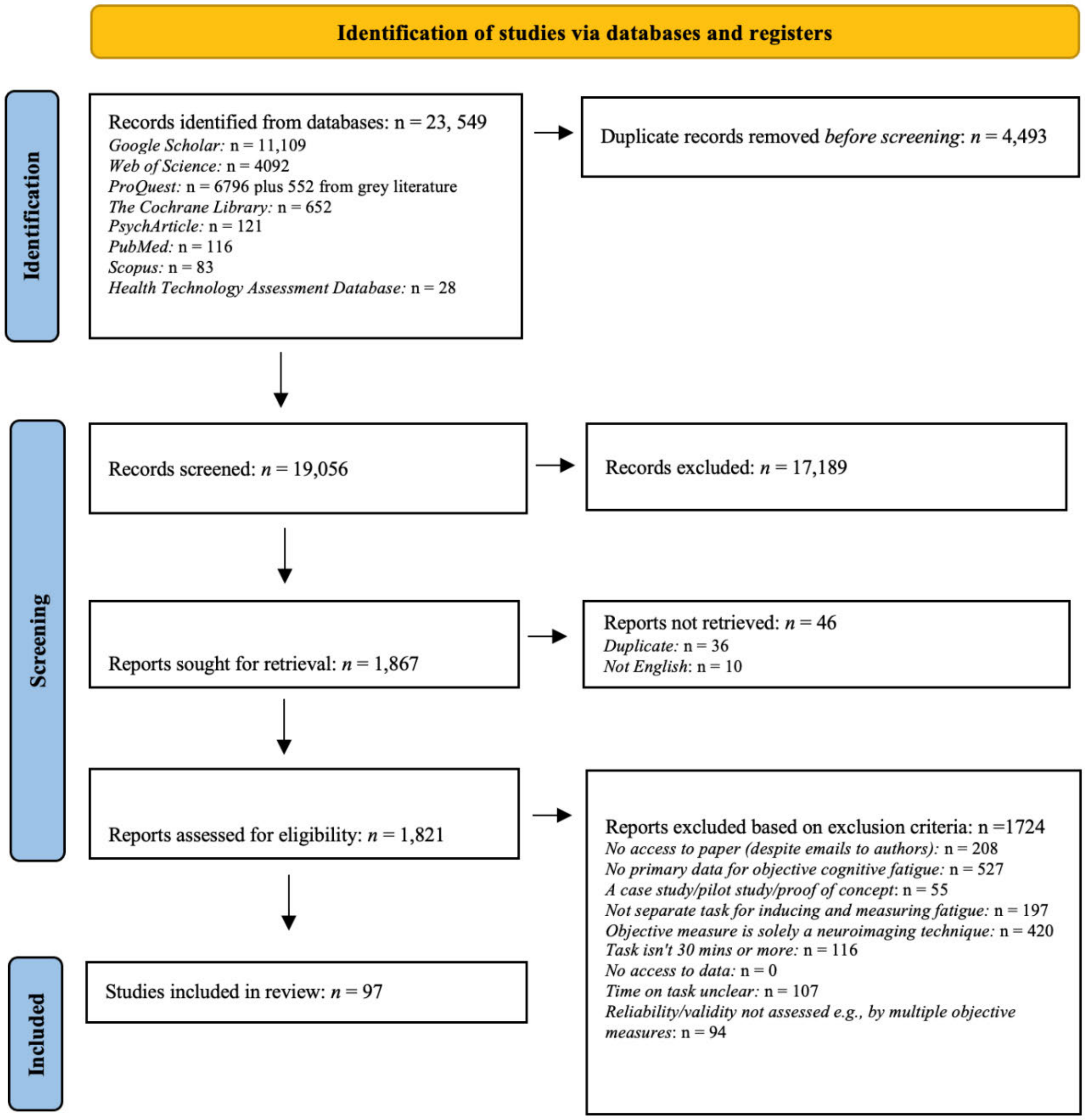
However, due to the large volume of studies that were still included after this stage, and the need for a reliable objective measure of cognitive fatigue, the following additional exclusion criteria were added:

1. No access to the study data;
2. The objective measure of cognitive fatigue was solely a neuroimaging technique;
3. The time-on-task used to induce cognitive fatigue was unclear;
4. Reliability/validity was not assessed by using more than one measure of objective cognitive fatigue.

A third reviewer (AC) resolved any conflicts.
The PRISMA diagram illustrates the process by which studies were excluded for final analysis.

Figure 2

The PRIMSA diagram showing the process of identification, screening and exclusion criteria.



Source: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Data Extraction

All articles deemed eligible underwent a manual search of their references to identify any additional studies that may meet the eligibility criteria, resulting in the finding of 23 further studies. Subsequently, the references of these 23 studies were also manually searched, yielding two more eligible articles. No additional studies were discovered in the reference search of these two studies. A total of 128 studies underwent data extraction, with two reviewers (AD and KS) independently extracting data from all 128 studies. To ensure accuracy, a second reviewer (AC) independently verified 10% of these extractions. Extracted data encompassed details such as the method of inducing cognitive fatigue, duration of induction, objectively measures of cognitive fatigue, and the relevant descriptive and inferential statistics. Further information extracted included the study title, authors, publication year, sample characteristics, overall study design, aims and hypotheses, conclusions, and any strengths and limitations.

During data extraction, two further duplicates were identified and removed. One study was excluded due to the population being under 18. Additionally, a further 50 articles were excluded for failing to meet criteria, not having a baseline comparison, or having incomplete data analysis. This resulted in a total of 76 articles that were subsequently assessed for risk of bias.

Risk of Bias

The assessment of bias risk was determined using the Joanna Briggs Institute Critical Appraisal questionnaire (Aromataris et al., 2020; Joanna Briggs Institute, 2014). Due to the types of studies sourced for this review, the quasi-experimental checklist was decided to be the most appropriate tool for risk of bias assessment, and appraised studies as ‘include’, ‘exclude’ or ‘seek further information’.

Due to the aims of this systematic review, reliability of the studies was considered to be the most important factor within the risk of bias tool. The checklist utilised two questions out of the nine total items to assess reliability: “Were outcomes measured in a reliable way?” and “Was appropriate statistical analysis used?”. The first of these two questions were assessed based on the reliability criteria already implemented, in which studies required two objective measures of cognitive fatigue and needed a baseline comparison. However, the second question regarding appropriate statistical analysis was more challenging to asses due to the limited number of studies that provided complete statistical transparency regarding power analysis. In order to fairly assess studies that did not provide this information, an *a priori* power threshold was determined. This meant the overall appraisal from the risk of bias tool did not immediately determine these studies to be ‘exclude’ or ‘seek further information’.

Table 3 outlines the *a priori* power analysis, determined using GPower, for the most commonly used statistical tests within the studies. This was to provide a guideline for *a priori* power analysis where information regarding power was not disclosed in the individual studies. Based on these guidelines, studies were deemed to be underpowered, and therefore, unreliable.

Table 3

Simulated a priori *power analysis guidelines from GPower based on analysis of commonly used statistical tests.*

Statistical Test	Minimum guideline sample size
t test	15
One way ANOVA (with 3 levels)	10
2 x 2 ANOVA	12
2 x 2 x 2 ANOVA	16

Among the assessed papers, twenty-four studies were considered underpowered and, consequently deemed not statistically reliable. As a result, these 24 studies were excluded from the narrative synthesis.

The final narrative synthesis incorporated findings from 56 studies derived from 52 journal articles.

RESULTS

Sample Characteristics

Of the 56 studies, 52 studies included healthy individuals only, covering a total population of 2853 participants (see *Table 4*). Four studies included participants with a health condition (Claros-Salinas et al., 2010; Hopstaken et al., 2015; Jonasson et al., 2018; Riese, 1999), totalling 214 participants. Whilst most studies recruited participants from the general population, Naeeri and colleagues (2021) and LeDuc and colleagues (2005) recruited only pilots, Feyer and colleagues (1994) and Williamson and colleagues (1996) recruited only professional truck drivers, Fortes and colleagues (2021) recruited only amateur boxers and Habay and colleagues (2021) recruited only experienced table tennis players.

Eight studies used a male-only population (Balasubramanian & Bhardwaj, 2018; Jagannath & Balasubramanian, 2014; Ma et al., 2018; Riemersma et al., 1978*b*; Riese, 1999; Saito et al., 2018; Zhang et al., 2008; Zhang & Yu, 2010). Eight studies did not report the sex of their participants (Feyer et al., 1994; Ioannucci, 2022*d*; LeDuc et al., 2005; Mehmood et al., 2022; Naeeri et al., 2021; Pattyn et al., 2008; Riemersma et al., 1978*a*; Williamson et al., 1996).

Table 4

Study sample characteristics

Study	Population	Sample size	Gender/Sex	Mean age (range)
Riemersma et al. (1978a)	University students	12	N/A	20
Riemersma et al. (1978b)	University students	12	M = 12, F = 0	25
Feyer et al. (1994)	Professional truck drivers	37	N/A	37.4
Williamson et al. (1996)	Professional truck drivers	27	N/A	38.4
Riese et al. (1999)	Closed head injury patients	16	M = 16, F = 0	26
Li et al. (2004)	Drivers	40	M = 20, F = 20	27.4 (18-30)
LeDuc et al. (2005)	U.S. Army Apache aviators	53	N/A	32
Liu & Wu (2008)	Drivers	24	M = 12, F = 12	N/A (24-30)
Pattyn et al. (2008)	University students	21	N/A	N/A
Zhang et al. (2008)	Male graduate students	14	M = 14, F = 0	23 (21-26)
Claros-Salinas et al. (2010)	MS and Stroke patients	42	M = 57, F = 61	40.9 (20-60)
Massar et al. (2010)	Students	12	M = 9, F = 3	22.2 (19-28)
Zhang & Yu (2010)	N/A	18	M = 18, F = 0	23.1 (21-26)
McKinley et al. (2011)	N/A	10	M = 9, F = 1	29 (18-42)
Tanaka et al. (2011)	N/A	19	M = 4, F = 15	43.6
Jagannath & Balasubramanian (2014)	Male drivers	20	M = 20, F = 0	23
Gergelyfi et al., (2015)	Sudoku players (non-smokers)	18	M = 10, F = 8	25.72 (20-35)
Guo et al. (2015)	Undergraduate students	36	M = 18, F = 18	20.3 (18-22)
Hopstaken et al. (2015)	Dysphoric and healthy students	82	M = 19, F = 63	21.1
Körber et al. (2015)	N/A	20	M = 18, F = 2	23.3
Hopstaken et al. (2016)	University students	47	M = 15, F = 32	20.5 (18-25)
Völker et al. (2016)	Young & old clerical workers	32	M = 15, F = 17	43.11
Zhao et al. (2017)	University students	16	M = 13, F = 3	24.7 (21-27)
Balasubramanian & Bhardwaj (2018)	Male drivers (right handed)	35	M = 35, F = 0	22
Guo et al. (2018)	Drivers (university students)	60	M = 30, F = 30	23.25 (18-27)
Jonasson et al. (2018)	Acquired Brain Injury	74	M = 31, F = 43	41.9
Ma et al. (2018)	Male drivers	21	M = 21, F = 0	40.1 (29-47)
Otto et al. (2018)	University students	16	M = 5, F = 11	24
Saito et al. (2018)	Male non-smokers	17	M = 17, F = 0	33.8
Wang et al. (2018)	N/A	12	M = 10, F = 2	28
Zargari et al. (2018)	Young and elderly groups	38	M = 18, F = 20	40.5
Herlambang et al. (2019)	University students	20	M = 8, F = 12	24.95
Smith et al. (2019)	Non-smokers	17	M = 11, F = 6	21.4 (18-27)
Van Cutsem et al. (2019)	Badminton players	20	M = 10, F = 10	24
Ferreira et al. (2020)	N/A	20	M = 10, F = 10	25.6 (18-35)
Herlambang et al. (2020)	University students	28	M = 17, F = 11	24.57
Wilson et al. (2020)	Sleep deprived participants	18	M = 16, F = 2	21.8
Chen et al. (2021)	Right handed adults	34	M = 19, F = 15	23.94 (22-29)
Fortes et al. (2021)	Amateur boxers (non-smokers)	21	M = 13, F = 8	23.33
Habay et al. (2021)	Experienced table tennis players	11	M = 7, F = 4	24
Hu & Lodewijks (2021a)	Drivers	20	M = 12, F = 8	26.4 (19-37)
Hu & Lodewijks (2021b)	N/A	40	M = 24, F = 16	26.2
Magnuson et al. (2021)	N/A	20	M = 8, F = 12	21-31
Naeeri et al. (2021)	Pilots	20	N/A	N/A

Study	Population	Sample size	Gender/Sex	Mean age (range)
Qin et al. (2021)	University students	11	M = 8, F = 3	22 (22-26)
Guillemin et al. (2022)	MS patients	38	M = 10, F = 28	31.5
Ioannucci (2022a)	N/A	48	M = 24, F = 24	22.1
Ioannucci (2022b)	N/A	14	M = 1, F = 13	21.3
Ioannucci (2022c)	N/A	16	M = 7, F = 9	21.9
Ioannucci (2022d)	N/A	24	N/A	21.6
Lampe & Deml (2022)	N/A	35	M = 21, F = 14	32.1 (19-61)
Mehmood et al. (2022)	Construction equipment operators	16	N/A	32.63
Shi et al. (2022)	Postgraduate students who drive	32	M = 20, F = 12	25.5
Sun et al. (2022)	N/A	17	M = 9, F = 8	N/A (25-45)
Yu et al. (2022)	N/A	35	M = 16, F = 19	26.4 (18-40)
Zhao et al. (2022)	Student novice drivers	58	M = 29, F = 29	22.03 (20-30)

Study Characteristics

The cognitively fatiguing tasks ranged from 30 minutes to 480 minutes. Most studies induced cognitive fatigue for 30 to 150 minutes. Thirteen studies used a cognitive fatigue task for 60 minutes, making this the modal time-on-task.

A variety of contexts were used to investigate cognitive fatigue. Whilst some studies investigated naturalistic scenarios whilst participants were at work (n = 17), others simulated these scenarios within simulators (n = 18), and 15 studies used an artificial computer-based task specifically designed to induce fatigue (such as the n-back task) before measuring performance on another task.

There were a variety of stand-alone induction tasks used by the remaining studies, which will each be discussed in turn.

A total of 18 studies included an induction task that has also been used for cognitive fatigue measurement in other studies included in this review. This is not including driving or driving simulation, of which a total of 21 studies used driving as an induction task, with four studies using driving performance as a measure. These studies will be disclosed in the task crossover section below. Table 5 provides an overview of study characteristics, the different types of procedure used to induce cognitive fatigue and the different objective measures of cognitive fatigue.

Table 5

An overview of all study characteristics, including the induction and measurement techniques for objective cognitive fatigue

Study	Setting	Induction task		Objective measures	Fatigue present	
		Task	Duration		Objective	Subjective
Riemersma et al. (1978a)	Driving	Driving	480	HR; intermittent reporting task (report kms travelled); colour changing task	Yes	N/A
Riemersma et al. (1978b)	Driving	Driving	48	HR frequency; HR variability; auditory choice reaction time task	Unclear	N/A
Feyer et al. (1994)	Driving (real world)	Driving	150	HR; episodic secondary task; critical flicker fusion frequency; vigilance task; unstable tracking	Yes for some measures	Unclear
Williamson et al. (1996)	Driving	Driving	149	HR; vigilance test; critical flicker fusion frequency; unstable tracking	Yes for some measures	Yes
Riese et al. (1999)	Driving simulator	Driving simulation	50	HR; selective attention task; BP; priming and recognition task; memory search task; dot counting task; peripheral stimuli	Yes for some measures	Yes
Li et al. (2004)	Driving on a test course	High speed driving	180	Vigilance task (reaction times); HR variability; critical flicker fusion frequency	Yes for some measures	Unclear
LeDuc et al. (2005)	Helicopter flying	Helicopter flight	72	Pupil diameter; pupil constriction; pupil latency; pupil constriction amplitude; saccadic velocity	Yes for some measures	Yes
Liu & Wu (2008)	Driving simulator	Driving simulation	60	Arithmetic task performance; distance estimation task	Yes for some measures	Yes
Pattyn et al. (2008)	Laboratory	Cognitive vigilance task	90	Respiratory inductive plethysmography; abdominal and ribcage excursions	Yes for some measures	Unclear
Zhang et al. (2008)	Laboratory	Switch task	120	EEG; HR variability	Yes for some measures	Yes
Claros-Salinas et al. (2010)	Laboratory	Normal daily routine	60	Go/No-go task; divided attention task; simple reaction time task	Yes for some measures	Yes
Massar et al. (2010)	Driving simulator	n-back task	90	Driving performance; EEG; auditory oddball task	Yes for some measures	Yes
Zhang & Yu (2010)	Laboratory	Mental arithmetic task	120	HR; EEG	Yes	Yes
McKinley et al. (2011)	Air force-relevant environment	Flying an unmanned aerial vehicle, target acquisition task	30	Psychomotor visual task; aircraft landing task; percentage of eye closure; approximate entropy for pupil position	Yes for some measures	N/A
Tanaka et al. (2011)	Laboratory	n-back task	30	Advanced trail making test; ECG	Yes for some measures	Yes
Jagannath & Balasubramanian (2014)	Driving simulator	Driving simulation	60	HR; surface electromyography; EEG; seat interface pressure; BP; oxygen saturation level	Yes for some measures	Unclear
Gergelyfi et al., (2015)	Laboratory	Sudoku	120	EEG; ECG-HR variability; pupil diameter; eye blink frequency; skin conductance response; working memory task; eye movements	Yes for some measures	Yes
Guo et al. (2015)	Generic workplace/ daily life	AX-continuous performance test	60	Go/No-go task; EEG	Yes for some measures	Yes
Hopstaken et al. (2015)	Laboratory	n-back task	60	Pupil diameter; stimulus-evoked pupil dilation	Yes	Yes
Körber et al. (2015)	Driving simulator	Driving simulation	42.5	Vigilance task; blink frequency; blink duration; pupil diameter	Yes for some measures	N/A

Study	Setting	Induction task		Objective measures	Fatigue present	
Hopstaken et al. (2016)	Laboratory	n-back task	90	Gaze position; pupil diameter; EEG	Yes for some measures	Yes
Völker et al. (2016)	Workplace	Working day (clerical)	480	Psychomotor speed; number of characters typed; typing errors; eye blink frequency	Yes for some measures	Yes
Zhao et al. (2017)	Driving simulator	Driving simulation	90	EEG; oddball task; ECG; eye blink frequency	Yes	Yes
Balasubramanian & Bhardwaj (2018)	Driving simulator	Driving simulation	120	EEG, ECG	Yes	N/A
Guo et al. (2018)	Driving simulator	Driving simulation	90	Go/No-go task; EEG	Yes	Yes
Jonasson et al. (2018)	Laboratory	Reading comprehension task	30	Digit symbol coding; attentional blink test; computerized rest combining divided attention and working memory simultaneously	Yes for some measures	Unclear
Ma et al. (2018)	Driving simulator	Driving simulation	60	EEG; eye-tracking	Yes for some measures	Yes
Otto et al. (2018)	Aviation simulator	Helicopter cockpit training	240	MRI; n-back task	Yes	Yes
Saito et al. (2018)	Laboratory	n-back task	40	Simple colour reaction task; Stroop task; accelerated plethysmography;	No	Yes
Wang et al. (2018)	Driving simulator	Driving simulation	240	ECG; EEG	Yes	Unclear
Zargari et al. (2018)	Laboratory	Pattern recognition task	40	Eye blink duration; eye blink frequency; saccade duration; peak velocity of eye movement; pupil dilation range	Yes	Yes
Herlambang et al. (2019)	Laboratory	Vowel counting pseudoword task	150	Pupil diameter; gaze positions; saccade frequency; eye blink frequency; ECG	Yes for some measures	Yes for some measures
Smith et al. (2019)	Laboratory	3 conditions (within subjects on different days); AX-Continuous Performance Test, Stroop and Psychomotor vigilance task, documentary watching (control)	45	EEG; HR variability; psychomotor vigilance test	Yes for some measures	Yes
Van Cutsem et al. (2019)	Laboratory	Stroop or documentary watching (control)	90	Visuomotor task; flanker task; HR	Yes for some measures	Yes
Ferreira at al. (2020)	Laboratory	Stroop or control condition	30	HR variability; reaction time task	Yes	Yes
Herlambang et al. (2020)	Laboratory	Sudoku	39	Pupil diameter; eye blink frequency; gaze positions; saccade frequency; HR variability	Yes	Yes
Wilson et al. (2020)	Aviation simulator	Flight simulation (take off and climb phase, then cruise phase)	126	HR variability; photoplethysmogram	Yes	N/A
Chen et al. (2021)	Laboratory	n-back task	100	EEG; suspension task (group condition); maths task (group ok-0zcondition)	Yes for some measures	Yes
Fortes et al. (2021)	Sport/ laboratory	Coaching video or social media or video game	30	Stroop task; boxing decision-making task	Yes for some measures	Yes

Study	Setting	Induction task		Objective measures	Fatigue present	
Habay et al. (2021)	Laboratory	Stroop or documentary watching (control)	60	Flanker task; EEG; ECG; visuomotor performance	Yes for some measures	Yes for some measures
Hu & Lodewijks (2021a)	Driving simulator	Driving simulation	60	Pupil diameter; vigilance test (SDT); number calculation task; eye blink frequency; eye fixation; saccade frequency	Yes for some measures	Yes
Hu & Lodewijks (2021b)	Driving simulator	Driving simulation	60	Pupil diameter; vigilance test (SDT); number calculation task; eye blink frequency; eye fixation; saccade frequency	Yes for some measures	Yes
Magnuson et al. (2021)	Laboratory	n-back task	60	EEG; Go/No-go task; object hit and avoid task; pegboard task	Yes for some measures	Yes
Naeeri et al. (2021)	Aviation simulator	Flight simulation	60	Psychomotor vigilance task; eye fixation number; eye fixation duration; mean pupil dimeter; visual entropy	Yes for some measures	Yes
Qin et al. (2021)	Aviation simulator	Flight simulation	90	ECG; blink rate; pupil diameter; vigilance task; attacking task	Yes for some measures	Yes
Guillemin et al. (2022)	Laboratory	Time load dual back task	32	Blink frequency; pupil diameter	Yes for some measures	Yes
Ioannucci (2022a)	Laboratory	Auditory tasks (either: auditory n-back task, pitch-sequence reproduction task, or auditory switch task)	41	Pupil diameter; EEG; Texture Discrimination Task	Yes for some measures	Yes
Ioannucci (2022b)	Laboratory	Auditory tasks (either: auditory n-back task, pitch-sequence reproduction task, or auditory switch task)	41	Pupil diameter; Texture Discrimination Task	Yes	Yes
Ioannucci (2022c)	Laboratory	Auditory tasks (either: auditory n-back task, pitch-sequence reproduction task, or auditory switch task)	41	Pupil diameter; Texture Discrimination Task	Yes	N/A
Ioannucci (2022d)	Laboratory	Auditory tasks (either: auditory n-back task, pitch-sequence reproduction task, or auditory switch task)	40	Texture Discrimination Task; fMRI	Yes for some measures	Yes
Lampe & Deml (2022)	Driving simulator	Driving simulation	40	Eyelid distance; eye blink frequency; eye blink duration	Yes for some measures	Yes
Mehmood et al. (2022)	Workplace	Excavation task	60	Facial features; EEG	Yes for some measures	Yes
Shi et al. (2022)	Driving simulator	Driving simulation	90	HR variability; pupillary light reflex	Yes for some measures	Yes
Sun et al. (2022)	Workplace	Continuous performance test	120	HR variability; critical flicker fusion frequency; eye blink frequency; electrodermal current; pupil diameter; EEG; SART	Yes for some measures	Yes
Yu et al. (2022)	Laboratory	Go/No-go task	180	EEG; pupil diameter	Yes for some measures	Unclear
Zhao et al. (2022)	Driving simulator	Driving simulation	60	Mobile phone call task; pupil diameter; EEG	Yes for some measures	N/A

HR: heart rate; EEG: electroencephalogram; ECG: electrocardiogram; fMRI: functional magnetic resonance imaging; MRI: magnetic resonance imaging; SART: sustained attention to response task; SDT: signal detection task.

Cognitive Fatigue Inducing Procedures

Work Day

Four studies used literacy, reading, or a normal office-based work/study day to induce cognitive fatigue. Five studies used aviation, which required piloting an aircraft or flight simulator, (LeDuc et al., 2005; Naeeri et al., 2021; Otto et al., 2018; Qin et al., 2021; Wilson et al., 2020) and a further two studies used manual labour (McKinley et al., 2011; Mehmood et al., 2022) to induce cognitive fatigue. Although these tasks may not induce cognitive fatigue uniformly across all participants, it is the most realistic way to induce workplace cognitive fatigue.

Zargari and colleagues' (2018) implemented a functional computer task to induce cognitive fatigue in order to replicate prolonged computer work. This task required participants to memorise a pattern, retain this information in their working memory during a washout period, and then replicate the pattern using the computer mouse. Participants completed this process 240 times, resulting in a 40-minute time-on-task period. Moreover, the authors wanted to determine if there was a difference between the 20 young participants in comparison to the 18 elderly participants. Subjective ratings of cognitive fatigue were taken before and after completion of the functional computer task, as well as in the 12 intervals during the 240 cycles. Oculometrics were recorded throughout the task as an objective measure of cognitive fatigue. In terms of task performance, a large effect size was found for overall performance, which was higher in the young participants in comparison to the elderly participants ($\eta p^2 = 0.5$). The overall performance for both the elderly and young participants increased with time-on-task with a large effect size ($\eta p^2 = 0.1$). This suggests that the overall performance measurements had discovered a practice effect for this task, and a prolonged time-on-task would be required in order for overall performance to show the impact of cognitive fatigue. Contrastingly, subjective cognitive fatigue ratings, as determined by the Karolinska Sleepiness Scale, increased for both groups ($\eta p^2 = 0.2$), with no significant differences between the groups. This highlights the disparity between objective and subjective levels of cognitive fatigue, particularly where a practice effect is involved.

Jonasson and colleagues' (2018) study aimed to investigate cognitive fatigue using a test and retest design. This included a number of tasks from a neuropsychology test battery. In order to induce cognitive fatigue, participants undertook a reading comprehension test for 30 minutes. This was conducted both on participants who had suffered a brain injury (labelled the 'mental fatigue' group) and a control group of healthy participants. The control group performed significantly better on eleven out of the fourteen tests conducted in the test battery in comparison to the mental fatigue group, suggesting that reading comprehension induced, or exacerbated, cognitive fatigue, in the mental fatigue group. Furthermore, in four of the tests a significant time x group interaction was found; while the control groups' performance improved, the mental fatigue groups' performance declined. This ultimately suggests that reading comprehension, specifically with 30 minutes time-on-task, is not a sufficient time period to induce cognitive fatigue in healthy individuals. Similarly to Zargari and colleagues (2018), it appears that a practice effect can be seen in healthy individuals' performance, resulting in the overall performance increase within some of the cognitive tests.

Völker and colleagues' (2016) study required participants to complete a normal working day of clerical activities to induce cognitive fatigue. Participants' daily activities, overall work day length and the number of rest breaks were at their discretion, however, they were requested not to undertake any physically demanding work within the work rest breaks. Work days lasted between eight and nine hours. Objective mental tests pre- and post- participants work day was conducted, which showed an overall tendency of reaction time decreasing over time, suggesting an increase in cognitive fatigue as a result from the work day.

Claros-Salinas and colleagues' (2010) study looked at the effects of cognitive tasks that reflected participants' daily work routine. These cognitive tasks were given to participants to reflect their specific vocational backgrounds; therefore, tasks were individually selected and different across the participants. Cognitive performance and subjective ratings of cognitive fatigue were assessed at three time points over two consecutive days of completing these individualised cognitive tasks (morning, midday and afternoon) to determine if cognitive fatigue levels differed over time. Moreover, participants were either healthy, or had been diagnosed with multiple sclerosis or had previously had a stroke. Subjective cognitive fatigue was found to increase across day one time points and day two time points for all three group conditions. Stroke and multiple sclerosis patients clearly performed more poorly overall than control participants for all measures of cognitive performance (reaction time task, Go/No-go task and a dual-task paradigm), indicating higher levels of objective cognitive fatigue within these participants. Cognitive performance was found to decline in all group conditions throughout both days, apart from the reaction time task to assess alertness, whereby control participants' performance was similar throughout the time points over both days. Overall, these results indicate that the cognitive tasks provided to participants to reflect their work day were able to induce objective cognitive fatigue across all participants, as well as increasing feelings of subjective cognitive fatigue.

These studies collectively demonstrate that cognitive fatigue can be induced through prolonged and repetitive cognitive daily work tasks. It is likely that motivation plays a larger role in the subjective fatigue of the participants in this category, due to their personal feelings towards their workday tasks, resulting in a greater discrepancy between objective and subjective cognitive fatigue.

Manual Labour

Mehmood and colleagues (2022) and McKinley and colleagues (2011) both used manual labour for the induction of cognitive fatigue. Mehmood and colleagues' (2022) study uses participants who are in construction, as the focus of their study is safety for construction workers who are subjected to cognitive fatigue. Specifically, these participants were part of an excavation team. Their task was therefore to complete monotonous and prolonged excavation task for 1 hour, during which the NASA-TLX, video recordings of facial expressions and EEG readings were taken at four time intervals, including a baseline measurement.

Comparatively, McKinley and colleagues' (2011) study used a target acquisition and an unmanned aerial vehicle task to induce cognitive fatigue. These tasks were selected due to their relevance to ongoing real-world Air Force operations. In the target acquisition task, participants were required to identify potential enemy targets whilst navigating towards a pre-specified target within enemy territory. This task was completed for 30 minutes prior to the unmanned aerial vehicle task, which took a further 30 minutes to complete. The unmanned aerial vehicle task involved participants simulating the landing of an unmanned aerial vehicle task. This task was theorised to be the more challenging of the two, as it required manual flight control alongside a predetermined flight path using a ground control station with a limited field of view. A psychomotor vigilance task was also conducted for a further 10 minutes, however, the authors discuss this measure is relation to its sensitivity to sleep loss. Eye-tracking data and task performance data were collected throughout.

Aviation

The five studies which used aviation (LeDuc et al., 2005; Naeeri et al., 2021; Otto et al., 2018; Qin et al., 2021; Wilson et al., 2020) for cognitive fatigue induction were varied task-wise, and therefore need to be addressed individually.

Qin and colleagues' (2021) study used a flight simulator where participants manually operated an aircraft over a series of consecutive flight paths. The flight paths had various direction changes included in order to keep participants engaged. Otto and colleagues (2018) had participants engage in a helicopter cockpit training task to induce cognitive fatigue. This training session lasted for a total of four hours, with a short break in the middle. This was designed to induce a similar level of fatigue of a demanding work day. Further details on training specifics were not provided. Naeeri and colleagues (2021) also had participants perform a task for a four-hour duration, which comprised of four one-hour flights. This was a flight simulation whereby the pilot did not have visibility out of the window, meaning they need to use the information provided from flight instruments alone. LeDuc and colleagues (2005) recruited US Apache Aviator helicopter pilots in a training environment using both instructor and student pilots. Participants had a minimum of 72 minutes flight time, which occurred throughout different times of the day, to induce cognitive fatigue. Lastly, Wilson and colleagues (2020) used a simulated flight task that included a take-off phase, lasting 6 minutes, and a cruise phase spanning approximately 2 hours. The cruise phase was designed to maximise the duration of low cognitive workload to steadily increase cognitive fatigue over time.

MATHS TASKS

Similarly to implementing reading or work day related tasks or allowing participants to carry out their usual work tasks as a method of cognitive fatigue induction, maths tasks were also used in three studies to research the resultant cognitive fatigue induction. Two studies had participants complete sudoku puzzles for extended time periods, and one study used a mental arithmetic task.

The mental arithmetic task used in Zhang and Yu's (2010) study assessed whether heart rate and EEG could be used to estimate objective cognitive fatigue levels. The task itself required participants to solve randomly generated sums that included four single digits and three operator signs, which were either a plus or minus sign. They then had to indicate whether the answer was greater than, equal to, or the less than a given digit, and respond by pressing the appropriate key (<, =, or >) . The task did not move on until participants had responded, but they were instructed to answer as quickly as possible. This task was undertaken until two hours had passed or participants quit from volitional exhaustion, with their task performance (reaction time and error rate) recorded throughout. Response time increased significantly between pre- and post- test measurements, and a non-significant increase in error was found. Subjectively, cognitive fatigue levels rose significantly. The authors concluded that these results indicated an increase in cognitive fatigue levels due to the prolonged time completing the arithmetic task.

Both Gergelyfi and colleagues (2015) and Herlambang and colleagues (2020) used sudoku as a fatigue inducing task. Both studies also investigated the role of motivation alongside cognitive fatigue. Herlambang and colleagues (2020) had participants complete computerised sudoku puzzles continuously for three hours without a break. This consisted of 14 blocks of alternating conditions: low levels of intrinsic motivation (presented in green) and high levels of intrinsic motivation (presented in black). Comparatively, Gergelyfi and colleagues (2015) had participants undertake sudoku puzzles for four blocks of 30 minutes, with a cognitive fatigue evaluation taking place between each block. The authors also chose to individualise the difficulty level of the sudoku puzzles for each participant, whereby the threshold for each participant was the ability to complete a sudoku grid with a median reaction time between five and ten seconds for input of individual cells, and less than five errors for the grid overall. This decision was made based on a pilot study that indicated higher difficulty levels resulted in higher median reaction times, and therefore it was necessary to alter the difficulty to a comparative threshold for all subjects to reduce limitations of this study. Subjective ratings indicated a significant increase in cognitive fatigue following the sudoku task (p = 0.0001), as well as changes found in physiological measures that indicated an increase of objective cognitive fatigue. Herlambang and colleagues (2020) also found changes in physiological measures. Time-on-task was a prominent factor throughout both high and low levels of intrinsic motivation, with more pronounced decline in physiological results being revealed in the sudoku puzzles with low levels of intrinsic motivation. Similarly, performance measures declined over time, but more significantly in the low-level condition. Both studies are therefore indicative of sudoku puzzles being able to cause both subjective and objective levels of cognitive fatigue due to utilising prolonged time-on-task.

Overall, all three studies strongly suggest that maths tasks can be used to induce cognitive fatigue, provided that the time-on-task is long enough to use significant levels of cognitive resources.

AX-CPT

Four studies used the AX-continuous performance task, which involves participants responding to the letter stimuli 'X' after being presented with a cue letter 'A'. They also have to respond to non-target stimuli (letters 'B' or 'Y') when presented in particular combinations with 'A' or 'X'. A further study (Sun et al., 2022) modified the continuous performance task by using numerical stimuli rather than letters.

Each of these studies confirmed that the task had caused fatigue by assessing performance indicators for the task itself (e.g., response times and error rates) and by using a subjective fatigue scale before investigating separate, objective measures. Guo et al. (2015) used the visual analogue scale to evaluate fatigue severity (VAS-F) after 90 minutes of their AX-continuous performance task tasks (finding large to very large effect sizes representing the difference between pre and post measures). They then used vigilance task performance, the critical flicker fusion task and heart rate variability as their objective measures of fatigue. Similarly, Smith and colleagues (2019) also used the visual analogue scale, as well as the Brunel Mood Scale, as subjective measures following a 45-minute treatment of either the Stoop, a psychomotor vigilance task or the AX-CPT. The AX-CPT resulted in a significant increase in subjective cognitive fatigue in both the visual analogue scale and the Brunel Mood Scale. A short, 3-minute psychomotor vigilance task, EEG and heart rate variability were used as objective measures following one of the three 45-mintue treatments.

In comparison, Chen et al. (2021) used the Stanford Sleepiness Scale to monitor fatigue induction for 100 minutes. Rather than comparing post scores to baseline scores, for the subjective measure, they compared the scores of participants who performed the task to a control group that did not perform the task. They found scores significantly higher in the fatigued group in comparison to the control group. They then investigated cognitive fatigue objectively using a maths task and suspension task performance. Sun et al. (2022) also confirmed fatigue using a sleepiness scale (the Karolinska Sleepiness Scale) rather than one specifically designed to measure subjective fatigue. They used a 120-minute continuous performance task to simulate working conditions for visual display terminal operators in coal mines. The analysis of the subjective measure indicated that the continuous performance task had successfully induced cognitive fatigue, and this was further evidenced in the objective measures of reaction time, error rate, pupil diameter and blink frequency and other physiological data.

These studies indicate that the AX-CPT is capable of strongly inducing subjective cognitive fatigue, but robust induction of objective cognitive fatigue seems to be dependent on prolonged time-on-task.

SWITCH TASK

The only paper that used a switch task for cognitive fatigue induction was Zhang and colleagues' (2008) study. The switch task is when a stimulus is presented in a two-by-two display square, with each trial moving the stimulus clockwise once around the two-by-two boxes. In this study, the stimuli were a randomised digit, in either blue or red, from zero to nine. Participants were required to use a left or right mouse click to respond to the colour if the stimulus was presented in the top two squares, or use the mouse click to respond to whether the number was odd or even when presented in the bottom two squares. The continuous switching between the response requirements every two trials as well as performing quickly with accurate responses for a prolonged time can lead to cognitive fatigue. In this study, participants performed the switch task for a minimum of two hours or until the participant quit due to volitional exhaustion. Subjective scores for cognitive fatigue significantly increased post-task, as well as percentage error rate. However, reaction times were almost identical pre- and post-task. Physiological measures showed a significant difference between pre- and post-task measurements, indicating that the switch task was able to induce cognitive fatigue over this time frame.

To conclude this section regarding induction of objective cognitive fatigue, the primary outcome is that time-on-task needs to be sufficient. Although our inclusion criteria only required time-on-task to be upwards of 30 minutes, it is likely that some tasks will fatigue individuals quicker than others. It is therefore important to consider this when using any of these tasks in future research, and completing tasks for a minimum, prolonged length of time or until volitional exhaustion may be a more robust and reliable way to utilise these tasks.

INDUCTION AND MEASUREMENT TASKS

The next section of the results summarises tasks that were for both the induction and measurement of cognitive fatigue. No papers discussed used the same task for induction and measurement within the same study.

STROOP TASK

The Stroop task is commonly used to induce cognitive fatigue as it is an indicator of inhibition ability. As the literature indicates that cognitive fatigue is one of the most predominant factors in decreased response inhibition (Guo et al., 2018), it is often assumed that the decrease in resources resulting in a decreased response inhibition (tracked over time using task performance measures such as response time and error rates) should correlate with an increase in cognitive fatigue levels over time.

Three studies used the Stroop task for cognitive fatigue induction (Ferreira et al., 2020; Habay et al., 2021; Van Cutsem et al., 2019). Habay and colleagues (2021) additionally recorded task performance of the 60-minute Stroop task while inducing cognitive fatigue and found that, although subjective measures of cognitive fatigue increased in comparison to the control group, there were no indicators of cognitive fatigue increasing within the task performance of the Stroop itself. Similarly, Ferreira et al., (2020), found an increase in both perceived cognitive fatigue, as well as reaction time to a secondary visuomotor task, for up to 15 minutes after participants completed a 30-minute Stroop task to induce fatigue. However, no significant differences in heart rate variability were detected between the cognitive fatigue group and the control group, which raises discussion points surrounding heart rate as a reliable indicator of cognitive fatigue increase. Van Cutsem and colleagues' (2019) study continues to evidence both Ferreira and colleagues' (2020) and Habay and colleagues' (2021) studies; the 30-minute Stroop task resulted in an increased level of perceived fatigue, with a secondary visuomotor task revealing increased reaction times. However, reaction times within the Stroop task itself actually decreased over time independently of the condition and control group. This could be indicative of participants experiencing the practice effect during the Stroop task, whereby task performance increases due to practice before plateauing. The lack of significant differences seen in reaction times and task performance of the Stroop throughout these studies may be due to time-on-task ultimately not being long enough to see these effects.

Two studies used the Stroop task as a measure of cognitive fatigue (Saito et al., 2018; Fortes et al., 2021). Fortes and colleagues (2021) used a 90 second Stroop task pre and post cognitive fatigue induction and found no changes in accuracy over condition or time. However, a moderate effect size ($\eta p^2 = 0.03$) for condition and a medium effect size ($\eta p^2 = 0.07$) of time effect on response time was found. A moderate effect size ($\eta p^2 = 0.05$) was also found for condition x time interaction, with the control, and cognitive fatigue induction tasks of smartphone use, and videogame use, all showing an increase in reaction times. The videogame and smartphone conditions also resulted in a higher response time in comparison to the control group, indicating that the Stroop task was able to distinguish between cognitive fatigue levels, but not tasks that require similar resources.

Comparatively, Saito and colleagues' (2018) study explored the effect of odorants on fatigue and used a novel modified Stroop task for six minutes, in which the word (a Japanese character) did not match the presented traffic light colour following a 40-minute n-back task to induce cognitive fatigue. A simple colour reaction task was also utilised; however, this was simply for confirmation that participants could distinguish and report colour changes correctly. The results revealed a significant increase in the percentage correct response rate, but not of reaction times, of the Stroop with odorant exposure in comparison to without odorant exposure. This suggests that the odorant exposure has an objectively anti-fatiguing effect, although no significant difference was found between the odorant exposure and non-exposure for subjective cognitive fatigue scale, evidencing the discrepancy between objective and subjective cognitive fatigue.

Fortes et al. (2021) found that the Stroop task revealed increased reaction times following cognitive fatigue induction via smartphone and video game use, though accuracy remained unchanged. Saito et al. (2018) showed that odorant exposure improved correct responses in a modified Stroop task following cognitive fatigue, despite no significant difference in subjective fatigue ratings. This suggests that the Stroop task could be utilised to differentiate between fluctuations for both increasing and decreasing levels of objective cognitive fatigue.

GO/NOGO TASKS

The Go/NoGo task is similar to the Stroop task in terms of the theory behind inhibition and cognitive resources. Participants are presented with a number of target and non-target stimuli which they must respond to or inhibit their response to accordingly. One study used a Go/NoGo paradigm as an inducer (Yu et al., 2022), with three studies using a Go/NoGo paradigm for a cognitive fatigue measure (Guo et al., 2018; Guo et al., 2015; Claros-Salinas et al., 2010). Yu and colleagues (2022) Go/NoGo trials included the use of letters and numbers presented at different angles (30° and 150°) for Go trials (i.e. R) and for NoGo trials with mirrored stimuli (i.e. Я). To help evidence the induction of cognitive fatigue this study recorded task performance for the Go/NoGo task. In this instance the NoGo trial is priority as this is the trial that tests the participants' ability to inhibit their response. Yu and colleagues (2022) found that stimuli presented at 30° had a significantly higher correct rejection rate than those presented at 150°, however, no main effect was found for task session. Interaction effects were found for session x rotation angle, and task session four x all prior task sessions in the 30° angle trials. These results suggest that the 150° angle was too difficult for participants to be able to respond with accuracy, thereby rendering the sensitivity of the 150° trials void in both conditions. However, the fourth and final task session showing significant differences to the previous three trials in the 30° trials is indicative of increased cognitive fatigue due to the increased inability to inhibit responses to the NoGo 30° trials.

Guo and colleagues (2015) used a Go/NoGo paradigm as measurement of cognitive fatigue for a control and music condition, of which music was expected to alleviate cognitive fatigue. They found that, although accuracy did not change, reaction times became significantly increased for the control, and did not change for the music condition. Comparatively in Guo and colleagues' (2018) study, where a 90-minute simulated driving task was used to induce cognitive fatigue, both reaction time and miss rate were significantly increased, but there was no significant difference shown between the driving and control group (who watched a movie instead) in the percentage of false hits on the NoGo trials either pre or post fatigue manipulation. Lastly, Claros-Salinas and colleagues' (2010) study used a 3-minute Go/NoGo session for measurement of cognitive fatigue in stroke and multiple sclerosis (MS) patients, as well as a healthy control group, at three time-points throughout the day for two consecutive days. Reaction times of the Go/NoGo increased for both the patient groups over both days, and decreased over the two days for the control group. In terms of error rates, no significant effects were found for group main effect or for group x time-of-day interaction effects. The Go/NoGo task therefore seems to be heavily dependent on the fatiguing task employed, as two of the three studies included in this review had no significant changes in accuracy, which are indicative of inhibitory response control.

n-Back

Eight studies used the n-back to induce cognitive fatigue (Hopstaken et al., 2016; Hopstaken et al., 2015; Ioannucci, 2022; Magnuson et al., 2021; Massar et al., 2010; Otto et al., 2015; Saito et al., 2018; Tanaka et al., 2011). The n-back is a task whereby participants must respond using a keyboard to signal when a stimulus is the same as the stimuli presented two stimuli previous. The n-back has been used to induce fatigue throughout literature and is the second most prominent technique for inducing cognitive fatigue in this review.

Hopstaken and colleagues have notably used the n-back in three studies included in our review. In their 2016 paper, the n-back was used to fatigue participants, as well as recording behavioural measures, including accuracy, (false alarm rate and hit rate) and reaction times. The authors note that the most relevant measure of performance was the accuracy of the n-back task, with time on task causing a significant decrease with a large effect size. However, in the final block of the n-back task whereby reward manipulation was implemented, the accuracy of participants in both studies increased significantly with a large effect size. This suggests that reward manipulation and, therefore, motivation, can combat the effects of cognitive fatigue, which is an important subjective, situational factor to consider alongside any objective measure. In Hopstaken and colleagues' (2015) study, reward manipulation was not utilised. Instead, they compared a control group to a dysphoric (defined as ‘mildly depressed’) group. No significant difference was found between the control and dysphoric group for time-on-task. Accuracy within blocks one and two of the n-back showed stable task performance. The authors suggest that this is due to practice effects, whereby the learning of the task masks the onset of cognitive fatigue. Accuracy began to decline more rapidly in blocks three through six, with medium effect size. Lastly, reaction times did not significantly change throughout the six blocks, which the authors use as evidence against a speed-accuracy trade-off (i.e., the speed does not affect the accuracy, and therefore is not a confounding factor to accuracy being affected by cognitive fatigue induction). In contrast to their hypotheses, Hopstaken and colleagues (2015) did not find any group interaction for time-on-task, accuracy or reaction times, suggesting that objective cognitive fatigue was induced at a similar rate between the control and dysphoric group. The authors do note a significant difference between group interactions using subjective measures. This suggests that objective and subjective levels of cognitive fatigue can be misaligned in dysphoric individuals.

Tanaka and colleagues (2011) used the 2-back for 30 minutes to induce cognitive fatigue and subsequently used the advanced trail making task, ECG readings and the visual analogue scale after completion of the 2-back task to assess objective cognitive fatigue. The visual analogue scale scores and the total trial counts of the advanced trail making test did not significantly change between pre- and post- n-back measurement, regardless of condition. Error rate of the advanced trail making task also did not show any significant difference between pre- and post-cognitive fatigue induction, however there was a numerical trend of increased error rate post-n-back task. Advanced trail making task error rates also tended to increase subsequent to the n-back task, but not significantly.

Massar and colleagues' (2010) study used the n-back at two different levels (0-back and 2-back) for 90 minutes to induce cognitive fatigue, following a baseline driving simulation session and a post-fatigue driving simulation session. Interestingly, self-reported fatigue from the RSME showed a significant increase in subjective cognitive fatigue, however, objective driving simulation performance was not affected by either n-back condition. This would superficially suggest that the driving simulation did not induce objective cognitive fatigue. However, the P3a EEG wave that was analysed alongside the driving simulation task was reduced in the post-fatigue driving simulation, and furthermore, the reduction was more pronounced following the 2-back task than the 0-back task. Analysis showed that the difference was such that the reduction in the P3a was highly significant, in comparison to the 0-back, which did not have a significant effect. The P3a wave is an auditory event-related potential which occurs when an auditory stimulus focuses the participants' attention towards a deviant or novel sound (Parmentier, 2014). As the P3a wave is meant to increase following a novel auditory stimulus, the reduction in the P3a following the 2-back condition indicates a decrease in attention, indicating a decrease in cognitive resources. Therefore, in contrast to the objective driving simulation performance results, the P3a wave suggests that cognitive fatigue was induced by the 2-back condition.

Saito and colleagues (2018) studied participants' fatigue levels following a 40 minute 2-back task with or without odorant exposure. As previously discussed, the Stroop showed a significant decrease in correct responses following odorant exposure, but not reaction times. However, analysis and discussion surrounding the impact of the n-back task on the non-odorant exposure was not addressed separately to the comparative of the odorant exposure condition. Therefore, the impact of the 2-back task on the objective levels of cognitive fatigue are unclear.

Magnuson and colleagues' (2021) study used a 60-minute n-back to cognitively fatigue participants in order to explore fatigue recovery. EEG during the n-back showed significantly increased alpha power. This indicates a decrease in alertness and attention levels, suggesting a resultant increase in cognitive fatigue (Gharagozlou et al., 2015). Task performance of the n-back did not significantly change, however raw data shows an increase in reaction time and a decrease in response accuracy. The authors conclude that 60 minutes is not a long enough time-on-task period for the impairment of task performance, regardless of the changes in neural activity. However, these results indicate that task performance alone is not sensitive enough to register the initial increase in objective cognitive fatigue. Similarly, the subjective ratings that the authors collated alongside the n-back show a highly significant increase in subjective cognitive fatigue and task aversion, further evidencing the disparity between subjective and objective cognitive fatigue, where objective fatigue can be further differentiated into neural activity levels and task performance.

Of the eight studies that used the n-back, only two studies used the n-back solely to measure cognitive fatigue through task performance. Chen and colleagues (2021) used the n-back as the primary task, in order to be able to interrupt the task by either 'pausing' the n-back (suspension task) or redirecting participants to a secondary task (interruption task). This was to determine the effect of interruption on attention and working memory under a fatigued and non-fatigued state. The authors utilised a spatial n-back task, meaning that the position of squares in a grid were the stimuli, as opposed to letters, numbers or auditory targets. The AX-CPT was used for 100 minutes to induce cognitive fatigue in the fatigued state prior to running the n-back with an interruption condition. The authors found that both suspension and interruption led to differential EEG activity, suggesting attentional reallocation, task performance and reaction times were negatively affected. Behaviourally, interruptions decreased overall accuracy and increased reaction times of post-interruption trials of the n-back, with an exacerbated effect found in the fatigued versus the non-fatigued state, and in comparison, to the suspension task. For reaction times in the n-back, a significant main effect of fatigue state (fatigued and non-fatigued) and trial type (pre- and post- interruption or suspension) was found, with a large effect size for both ($\eta p^2 = 0.314$ and 0.768, respectively). Similarly, the main effects of fatigue state and trial type on n-back task accuracy were also found to be significant, with large effect sizes ($\eta p^2 = 0.275$ and 0.458, respectively). The large effect size despite the relatively small sample size indicates that it is a strong correlation. However, it is important to consider that the effect of the interruption may be confounded by the n-back increasing cognitive fatigue, as it is clear from the previously mentioned studies that the n-back is capable of doing this.

In the study by Otto and colleagues (2018), neural structures were identified using MRI following changes to fatigue and workload states. Participants completed the n-back task pre- and post- completion of a fatiguing helicopter simulation task. The n-back was presented with three different levels (1-back; 2-back; 3-back), with reaction times and error rates being measured, alongside a subjective rating scale of mental effort (RSME). A significant effect of task load (manipulated by the different n-back levels) on both error rates and the RSME was observed. However, reaction times did not show a significant effect with varying task load. In conjunction with the subjective RSME results showing only a slight increase for the 3-back task condition, in comparison to a disproportionate increase subjective mental effort in the easier 1-back task condition, the authors conclude that both the RSME and reaction time results reflect the monotony of the 1-back task. Consequently, the 3-back task was suggested to be more appealing for participants to actively engage in due to being more challenging than the 1-back task. This is an important factor to consider when analysing the MRI results. Although the 3-back task did not result in higher levels of subjective mental effort, and hitherto, subjective cognitive fatigue, the supplementary materials presenting the raw data for overall reaction time and mean errors shows a significant increase from the 1-back through to the 3-back. This strongly suggests that the 3-back requires more cognitive resources and is consequently more objectively cognitively fatiguing. This is evidenced further by the fMRI results, which indicate a load-related increase in neural activity in several brain regions as the n-back conditions increase. As a result, the n-back seems to be a difficult task to use for measuring cognitive fatigue objectively, as the higher load conditions further exacerbates objective cognitive fatigue in the process of the measurement, as shown by the raw reaction time and accuracy data, and the lower condition n-back task induces monotony, which results in an increase in subjective cognitive fatigue. This is a problematic factor for the use of this task in real-world terms for two main reasons: first, people will not wish to engage in the task due to boredom, resulting in disengagement from daily usage of a measure. Secondly, increasing either objective or subjective cognitive fatigue is not a desirable characteristic of the task.

A similar task used to induce cognitive fatigue was the TloadDback, which has alternating stimuli of letters and numbers. Participants must respond using different keys as to whether the number was odd or even as well as separately responding to whether the letter presented was the same as the letter previously presented. Guillemin and colleagues (2022) explored the difference of cognitive fatigue presentation between patients with multiple sclerosis and healthy control participants. They used a high cognitive load and a low cognitive load version of the TloadDback to induce two levels of cognitive fatigue. Accuracy and reaction time data was collected alongside subjective measures, as well as pupillometry measures during the TloadDback. Results showed that performance error rate and reaction times both increased as time-on-task increased in the high cognitive load condition, although a significant variance in reaction times was not found.

AUDITORY TASKS

Auditory tasks have been used both for the induction and measurement of cognitive fatigue in the papers included in this review. Ioannucci and colleagues (2022) used a combination of auditory tasks throughout their experiments to induce cognitive fatigue. Tasks were cycled for 41 minutes in total, before measuring cognitive fatigue with a behavioural index. Efficacy of the auditory tasks in the induction of cognitive fatigue were determined using a mixed model with pupillometry data, which was collected alongside the auditory tasks, and in conjunction with the two conditions of the auditory tasks, whereby participants completed either all the ‘easy’ or all the ‘hard’ versions of the tasks. The tasks included a modified n-back task. Participants were required to respond when a letter was the same as the letter presented two previously but using an auditory stimulus instead of a visual stimulus. Participants in the ‘easy’ group had an overall mean accuracy of 93.3%, in comparison to the 'hard' group, with an overall mean accuracy of 62.1% ($p < 0.0001$).

During the auditory pitch-sequence task, subjects replicated a sequence of auditory stimuli presented to them as beeps. Four different beeps were used in a randomly generated sequence, with one low-pitch, two medium-pitch and one high-pitch. Participants indicated the beep pattern with corresponding keyboard keys. For the easy condition group, only two beeps were used in a sequence, in comparison to the hard condition group which comprised of eight beeps. The hard condition group accuracy was 49.1% in comparison to the easy group, who reached 73.5% accuracy ($p < 0.0001$).

Lastly, an auditory switch task was utilised. This consisted of randomly presented sounds that occurred in either the right or left earphone. Stimuli were chosen from two categories of animal or vehicle sounds. The easy condition group simply had to respond to whether the sound was played through the left or right earphone. The hard condition group had a voice instructing which category of sounds they were required to respond to inversely at random intervals (i.e., if vehicle noises were specified, if the stimulus was presented in the right earphone, they would respond by pressing the left key and vice versa, while animal noises would remain as left earphone stimulus, left key pressing and vice versa). A further category of electronic sounds was also added, whereby participants were asked to not respond at all, adding a layer of response inhibition. Again, the easy condition group significantly outperformed the hard condition group, with an accuracy of 99.2% vs 89.1% respectively ($p = 0.0002$).

Auditory tasks were used as a measure in four studies (Feyer et al., 1994; Körber et al., 2015; Massar et al., 2010; Riemersma et al., 1978). Körber and colleagues (2015) used an auditory oddball paradigm during the driving task used to induce cognitive fatigue. Participants were instructed to react to high-pitched stimuli, which occurred for 20% of the trials, and ignore low-pitched stimuli which occurred for the remaining 80% of trials. Reaction times, misses and false hits were recorded in order to detect changes in vigilance. No significant effects were found for time-on-task for reaction time, missed targets or false hits.

Massar and colleagues' (2010) study also utilised an auditory oddball task, which was conducted during a simulated driving and non-driving condition. However, the participants were instructed to ignore all the tones presented. This task took place pre- and post- cognitive fatigue manipulation via the use of an n-back task. It is important to note that the authors used this task as a means to measure attention allocation, rather than cognitive fatigue. Arguably, attention allocation could be considered as a constituent of cognitive fatigue, however, as participants were instructed to ignore the tones presented to them, no data was collected specifically for oddball task performance. The simulated driving performance was analysed instead and was used as a measure of cognitive fatigue for pre- and post- n-back task completion.

Riemersma and colleagues (1978*b*) utilised an auditory choice reaction time task during their real-world driving task for induction of cognitive fatigue. They found a large difference in reaction times between the town and rural driving conditions, however, the authors describe their results as “disappointing”, as the expected mean reaction time differences were not revealed within the data. The authors suggest that this could be due to the participants performing this task, which was designed to be secondary, as a primary task, and therefore allocated too much attention to the task. They also note that technical difficulties were present in recording participants responses. For this reason, the auditory task in this study cannot reliably add towards the evidence base of using an auditory task for the detection of objective cognitive fatigue.

Lastly, Feyer and colleagues (1994) also used a secondary reaction time task during real-world driving, whereby an auditory stimulus was played and participants were required to give an oral response into a microphone (“yep”) as quickly as possible. These occurred 15 times at random intervals over the 30-minute driving task. Similarly to Riemersma and colleagues (1978*b*), no significant differences were found for driving trip type (single driver or two-person shared driving and short vs long trip) or time-on-task.

In summary, Körber et al. (2015) and Feyer et al. (1994) found no significant effects of time-on-task on reaction times, misses, or false hits during auditory tasks designed to measure vigilance in driving contexts. Massar et al. (2010) used an auditory oddball task to measure attention allocation rather than cognitive fatigue, and Riemersma et al. (1978) faced technical difficulties and found no reliable mean reaction time differences in their auditory task during real-world driving. Therefore, these studies do not provide a compelling argument to use an auditory task in terms of reliability, specifically within a driving context. Auditory tasks would need to be further investigated in other settings to further cultivate reliable data collection. This is an important avenue to explore in the future for individuals who are blind or have limited sight.

VISUAL TASKS

Various visual tasks were utilised throughout the studies included in this review. Of the selection, four tasks stood out in terms of usage for measurement: the texture discrimination task, the Flanker task, the critical flicker fusion task and unstable tracking.

Texture Discrimination Task

The texture discrimination task requires participants to identify a peripheral target that differs in orientation to the norm. For example, the texture discrimination task was used in all four of Ioannucci and colleagues' (2022) experiments and participants identified a left-tilted diagonal line that was presented horizontally (⟋) or vertically (\ ) against a background of lines that were horizontal (|). This was achieved by participants pressing set keys on the keyboard to record whether they had seen horizontal or vertical diagonal lines. The peripheral targets appeared in either the upper left or upper right quadrant of the background lines for each block.

For the first experiment, an inter-stimulus interval between 6 and 600ms was presented followed by a 100ms mask. This was incorporated to disrupt participants' processing, and consequently, trials with a shorter inter-stimulus intervals were expected to be harder. A complete texture discrimination task lasted a total of approximately 12 minutes, and consisted of four blocks, two per quadrant, which were randomised for every participant. Experiment one showed that shorter inter-stimulus intervals and increased session number (pre-test, mid-test, or post-test, where the test indicates a 41-minute visual saturation alongside auditory tasks designed to induce cognitive fatigue) resulted in a strongly significant decrease in percentage of correct responses.

The second experiment setup of the texture discrimination task was similar; however, the peripheral targets were confined to one quadrant as the participants were required to wear a patch over one eye for monocular viewing of the targets. The quadrant and eye were alternated per block for a total of four blocks. Individualised inter-stimulus intervals were used based on an 80% correct response rate threshold for each participant. Random inter-trial intervals were introduced to counteract the predictability of fixed inter-stimulus intervals. Similarly to experiment one, visual stimuli saturation lasted for 41 minutes between the pre- and post-task sessions, but was conducted on one eye only. The effects of session (pre-, mid- and post- task) and the effects of session and monocular viewing were both significant on the response accuracy for the texture discrimination task. Post-hoc testing revealed a significant decrease in performance accuracy which was found only in the saturated eye condition. Because the non-saturated eye condition when undergoing monocular viewing of the texture discrimination task did not incur a significant decrease in task performance, it can be suggested that the saturated eye was prone to visual fatigue. Therefore, this experiment seems to indicate that texture discrimination task performance can reflect visual fatigue change, but not necessarily overall objective cognitive fatigue change.

Experiment three was kept the same as experiment two in terms of the texture discrimination task setup, however, two-line orientations were used and a larger number of trials (128 per block) were incorporated in order to have an equal number of trials per combination of conditions to the previous experiments. The visual saturation was also kept the same as experiments one and two. Results showed a significant decrease in accuracy for the fixed effects of session condition and quadrant presentation, as well as a significant result for the interaction effect of session, quadrant and line orientation. Post hoc analysis revealed that the significant difference was primarily due to the difference between the pre- and post-test sessions.

Ioannucci and colleagues (2022) conducted four experiments using the texture discrimination task to measure the effects of cognitive fatigue, finding that shorter inter-stimulus intervals and increased session numbers significantly decreased correct response rates. Monocular viewing in the second and third experiments revealed that visual fatigue, particularly in the eye subjected to visual saturation, led to a significant decrease in task performance, highlighting the task's sensitivity to visual fatigue rather than overall cognitive fatigue.

The Flanker Task

During the Flanker task, participants are instructed to identify the orientation of an arrow as quickly and as accurately as possible while ignoring the two ‘flankers’, or distraction targets, on either side (e.g. < < > < <). This engages the need for inhibitory control, which requires high levels of cognitive resources and attention. The Flanker task was utilised by Habay and colleagues (2021) and Van Cutsem and colleagues (2019) for five and three minutes, respectively. Both studies collected the reaction time and task accuracy of the Flanker task pre- and post- their fatigue inducing task. The Flanker task had significant results for Van Cutsem and colleagues (2019) in terms of reaction time, but not for accuracy, whilst Habay and colleagues' (2021) study found no significant effect on either variable. Both these studies had similar cognitive fatigue induction protocols and used the Stroop task for induction and had a documentary watching group as a control condition. Therefore, the Flanker task needs to be more thoroughly researched in terms of its sensitivity to cognitive fatigue due to the contradicting results of these studies.

The Critical Flicker Fusion Task

The critical flicker fusion task was used in four studies as a means for measuring objective cognitive fatigue (Feyer et al., 1994; Li et al., 2004; Sun et al., 2022; Williamson et al., 1996) The level of cognitive fatigue can be determined by the threshold of individuals being able to see a singular light flash or the 'descending' threshold, as opposed to the frequency being high enough for the light flashing to merge together 'ascending' threshold (Mankowska et al., 2021).

Sun and colleagues (2022) define mild, moderate and severe fatigue as a decline in the threshold of 1.0 – 3.9Hz, 4.0 – 7.9Hz and greater than 8.0 Hz respectively. As this study is the most recent, the other results from the other three papers will be discussed individually and in line with these thresholds for comparative purposes where possible. Sun and colleagues (2022) found a significant mean decline in threshold of 4.49Hz following the induction of cognitive fatigue, indicating that participants had succumbed to a moderate level of fatigue ($p < 0.001$).

Feyer and colleagues (1994) also found that the descending threshold decreased over time-on-task during real-world driving, from a mean of 47.11 Hz pre-trip to 44.66 Hz post-trip for the single driver condition. This gives a total threshold decline of 2.45 Hz, which indicates mild cognitive fatigue induction according to Sun and colleagues' (2022) criteria. A significant effect of time ($p = 0.026$) and a significant linear trend in the data ($p = 0.007$) was found, indicating that the descending threshold was sensitive enough to identify mild objective cognitive fatigue induction. Li and colleagues' (2004) study also used real-world driving for fatigue induction and the critical flicker fusion task was undertaken before and after. The control group for this study (whereby participants undertook the driving task without the treatment of magnitopuncture) had a significant decline in critical flicker fusion threshold of 3.12Hz, equivalent to mild fatigue according to Sun and colleagues (2022).

Lastly, in Williamson and colleagues' (1996) study, the critical flicker fusion task provided mixed results following induction of cognitive fatigue via real-world driving. The pre-driving threshold had a mean of 38.5Hz, with the post-driving threshold increasing to a mean of 40.1Hz. This 1.6Hz increase in threshold implies that participants performance for the critical flicker fusion task improved and therefore, the driving task did not induce cognitive fatigue. A non-existent increase in cognitive fatigue during driving is, however, an unlikely conclusion. A potential reason for this result is that the study used an ascending threshold, rather than a descending threshold, and therefore the task was not sensitive to cognitive fatigue, however, this was not specified by the authors.

The descending critical flicker fusion task was used in three studies to measure cognitive fatigue, with Sun and colleagues (2022) defining mild, moderate, and severe fatigue based on specific declines in flicker fusion threshold. Sun and colleagues (2022) found a significant decline indicating moderate fatigue, while Feyer and colleagues (1994) and Li and colleagues (2004) observed declines consistent with mild fatigue. These results indicate that the critical flicker fusion task is sensitive to fluctuations in objective cognitive fatigue and has threshold traits that could be implemented as an adaptable, comparative measure of objective cognitive fatigue.

Unstable Tracking Task

Williamson and colleagues (1996) and Feyer and colleagues (1994) both used an unstable tracking task to measure objective cognitive fatigue as a result of real-world driving in their studies. The unstable tracking task evaluates the driver's proficiency in managing progressively unpredictable movements of a visual stimulus using a response dial. The variability of stimulus movement is directly correlated with the driver's performance capacity, i.e., as the driver's performance improves, the stimulus becomes more unpredictable. The assessment involves measuring both the duration the driver can maintain control over the stimulus and the level of stimulus instability, representing the task difficulty. Alertness can be linked to an extended period of control and higher difficulty levels. For Williamson and colleagues (1996), drivers showed a decline in performance around the middle of the trip duration. This was evident in both the time-on-task and task difficulty measures. However, participants in the single driver condition demonstrated an improvement from pre- to post-trip for time-on-task. Feyer and colleagues (1994) obtained similar results, with driver's unstable tracking performance remaining unchanged or decreasing around mid-trip and improving between pre- and post-trip measurements. These results indicate that the unstable tracking task was sensitive enough to differentiate between changes in drivers' performance, but, assuming that real-world driving is cognitively fatiguing, may not directly correlate to cognitive fatigue being the impacting factor on performance. Secondly, it challenges the concept of practice effects as performance initially decreases before increasing again.

Other visual tasks used for measurement of objective cognitive fatigue the non-driving performance measures undertaken by Riese and colleagues (1999), which included a memory search task, a selective attention task and a priming and recognition task. These tasks ran for five, four and five minutes respectively pre- and post the driving simulation task used to induce cognitive fatigue. All three of these tasks were assessed using reaction times, accuracy, heart rate and blood pressure. Descriptions of these experiments were not present in this paper, and there was no disclosure of the raw data collected from these tasks and consequently no inferential statistics were provided, with the exception of heart rate and blood pressure. However, the blood pressure results were non-significant, and were discussed under the term "physiological stress" and therefore have not been considered in line with objective cognitive fatigue. Similarly, heart rate variability had a non-significant decrease between pre- and post-test tasks and was discussed as a measure of "mental effort", which the authors assumed to be a logical variable of cognitive fatigue.

Driving and Driving Performance

Driving is a popular choice of task for inducing cognitive fatigue because it requires a high level of attention and concentration. It is also an accessible task that many participants can partake in with relative ease, with a high likelihood of being carried out on a regular basis within the realm of daily tasks. Driving is made more accessible as a cognitive fatigue induction technique due to the availability of driving simulators. Driving simulators are particularly useful in making experiments safe, meaning participants can be fatigued to a greater extent without the risk of road traffic accidents. It also allows for the experiment to be far more consistent, as each participant will use the same equipment, and can be subjected to exactly the same ‘road’ conditions if required (in terms of hazards and visibility) by using the same video. External environmental factors such as weather and temperature are also easily controlled and accounted for.

Confounding factors such as sleepiness, experience levels and road conditions (which are discussed in Johan et al., 1978, Di Flumeri et al., 2018; Ahn et al., 2016 and are evidenced to cause an increase in fatigue) can differ depending on individual differences and on a day-to-day basis. It is worth noting that driving can also increase stress levels and anxiety, quite often as a result of the aforementioned confounding factors (Abdullah & Mustapar, 2019), which can alter other, separate physiological measures of cognitive fatigue such as heart rate.

Driving as an induction task

Table 6 shows the five studies that used real-world driving, and, although this is a small number of studies, there is no explicit pattern as to whether real-world driving increases or decreases heart rate or heart rate variability significantly. Heart rate was measured in all five real-world driving experiments, although many of the inferential statistics were not reported. Fourteen studies implemented a driving simulator, but only three of them included heart rate metrics.

Table 6

An overview of all driving studies for cognitive fatigue induction, driving performance and secondary task performance, including heart rate results where applicable

	Study	Time-on-Task (mins)	Induction Task	HR Metric	Significant?	Comments	Increase/Decrease
Real World Driving							
	Riemersma et al. (1978a)	480	Driving (real-world; night time)	Heart rate	Yes		Decrease
	Riemersma et al. (1978b)	48	Driving (real-world; three short test routes during daytime: in town, on a four-lane motorway, and on a two-lane rural road)	Heart rate variability and frequency	Not Reported	Trend of increase in both	
				Heart rate variability	Not Reported	Trend of increase in variability	Increase
	Feyer et al., (1994)	150	Driving (real-world)	Heart rate	Yes	For single group only	
	Williamson et al., (1996)	149	Driving (real-world)	Heart rate	Not Reported		
	Li et al. (2003)	180	High speed driving on a test course	Heart rate variability	Not Reported		
Driving Simulator							
	Riese et al. (1999)	50	Simulated driving	Heart rate	No		N/A
	Liu & Wu (2008)	60	Simulated driving	-	-		-
	Jagannath & Balasubramanian (2014)	60	Simulated driving	Heart rate	Yes		Decrease
	Körber et al. (2015)	42.5	Simulated driving	-	-		-
	Zhao et al. (2017)	90	Simulated driving	-	-		-
	Balasubramanian & Bhardwaj (2018)	120	Simulated driving	ECG	Yes	Increase in HF only; LF and LF/HF ratio had decreasing trends but not significantly	Increase
	Guo et al. (2018)	90	Simulated driving	-	-		
	Ma et al. (2018)	60	Simulated driving	-	-		-
	Wang et al. (2018)	240	Simulated driving	-	-		-
	Hu & Lodewijks (2021a)	60	Simulated driving	-	-		-

	Study	Time-on-Task (mins)	Induction Task	HR Metric	Significant?	Comments	Increase/Decrease
	Hu & Lodewijks (2021b)	60	Simulated driving	-	-		-
	Lampe & Deml (2022)	40	Simulated driving	-	-		-
	Shi et al. (2022)	90	Simulated driving	Heart rate variability	Yes		Increase
	Zhao et al. (2022)	60	Simulated driving	-	-		-
Driving Performance Measures				**Driving Performance/ Secondary Tasks**			
	Riemersma et al. (1978a)	48	Driving (real-world; night time)	Lane Position	Yes	Standard deviation	Increase
				Speed maintenance	Yes	Standard deviation	Increase
				Intermittent speed reporting	Yes	Percentage of incorrect responses	Increase
				Colour changing task	Yes	Missed responses, incorrect responses and 'mental blocks'	Increase
	Riemersma et al. (1978b)	48	Driving (real-world; three short test routes during daytime: in town, on a four-lane motorway, and on a two-lane rural road)	Speed maintenance	Not Reported	Motorway condition	Increase
				Lane position	Not Reported	Motorway condition	Increase
				Mean amplitude of steering wheel deflection	Not Reported	Motorway condition	Decrease
	Riese et al. (1999)	50	Simulated driving	Lane position	Yes	Standard deviation of lateral position	Increase
				Dot counting	Yes	Secondary task; % of correct answers	Decrease
	Liu & Wu (2008)	60	Simulated driving	Lane Position	Yes	Standard deviation	Increase
				Lateral acceleration	Yes	Standard deviation	Increase
				Traffic sign distance estimation	Yes	Time judged in seconds	Decrease
	Massar et al. (2010)	90	n-back	Lane Position	No	Standard deviation of lane position	N/A

This is worth considering for a couple of reasons: first, that the authors' have an innate assumption that heart rate is more strongly affected by real-world driving than by a simulation, possibly due to awareness of confounding factors such as stress; and secondly, that previous literature may have found that simulators do not cause a reliable increase in heart rate (as would be expected from cognitive fatigue induction), and therefore, the authors' do not feel it is a useful metric to analyse. Nevertheless, based on heart rate metrics alone, neither real-world nor simulated driving are suggested to induce cognitive fatigue. In contrast, driving performance and secondary task performance provide a much stronger argument for cognitive fatigue induction form driving.

Driving Performance Measures

Reise and colleagues (1999) lane positioning measures included individualisation of difficulty. As driving was completed on a driving simulator, it was possible to manipulate side-wind as a factor. Side-wind was manipulated for a total of three minutes over the course of the driving task, in 12 blocks of 15 seconds each. Participants needed to stay within the lane boundaries for 90% of the time, if lane position deteriorated below 90%, the side-wind was decreased.

Secondary Tasks while Driving

Reise and colleagues (1999) included several secondary tasks (separately to their driving performance measure of lane positioning) while participants were driving in order to assess increase of cognitive fatigue level over time. This included divided attention task, requiring participants to complete a dot counting task and a peripheral stimuli task simultaneously. The dot counting task was presented in the middle of the simulator display, just below the horizon line, and required participants to indicate the number of dots presented (between five and nine) by pressing the correct number of a keypad. Task difficulty of the dot counting task was also individualised by altering the inter-stimulus interval determined by a pre-experimental training task. Participants who completed trials at a high speed without compromising accuracy had a smaller inter-stimuli interval during the experiment in comparison to participants with lower speed and accuracy. The peripheral stimuli task was presented on two monitors flanking the central simulator screen and required participants to detect traffics signs. These were represented as a simple triangle sign schematic with an arrow pointing either left, right, up or down. When the arrow was pointing up or down, participants were instructed to ignore the stimuli. When the arrow was pointing left or right, participants were instructed to press the horn located on the left or right side of the steering wheel respectively. Both the dot-counting task and the peripheral stimuli task performance were recorded in terms of accuracy and reaction times. Both tasks had significant decreases in response rate and accuracy, suggesting that attention rates decreased over time due to the driving simulation depleting cognitive resources. A five-minute memory search task, a four-minute selective attention task and a five minute priming and recognition task were completed pre- and post- fatigue induction, and were previously discussed fully in the visual task section.

The colour changing task used in Riemersma and colleagues' (1977) first experiment required participants to respond to a change in colour presented on a light mounted onto the dashboard of the car, by pressing the horn. Missed changes in colour, incorrect responses and mental blocks were calculated, where 'mental blocks' were defined as a reaction time exceeding twice the median. Both mental blocks and missed signals showed a significant increase with time-on-task. The kilometrage reporting task required participants to identify whenever their speed was a multiple of 20. The percentage of incorrect and missed reports were significantly increased from the first to the second part of the night time driving condition. The results from both secondary tasks indicate a decrease in attentional resources, thus indicating an increase in objective cognitive fatigue due to driving which seems to be exacerbated during nighttime driving.

OBJECTIVE COGNITIVE FATIGUE MEASURES

NEUROIMAGING

Neuroimaging can be used to monitor fatigue by assessing changes in brain function and activation. fMRI can be used to examine functional cerebral activity using blood oxygen level-dependent (BOLD) contrast, whereas positron emission topography (PET) scanning assesses the functional changes by using radioactive tracers to measure glucose metabolism (DeLuca et al., 2009). 15 studies included in this review used at least one form of neuroimaging to measure cognitive fatigue. The most popular neuroimaging technique within the included studies was EEG, which has been used to monitor cognitive fatigue through the examination of changes in the distribution, amplitude and frequency of alpha waves during fatigue onset. Alertness levels can be used as a correlate of cognitive fatigue level due to alertness being diminished through lack of cognitive resources or boredom (see *Figure 1*). It is noted that certain changes in alpha wave characteristics, as well as changes in relative energy levels of different energy bands, can signify alertness levels. For example, a decrease in (alpha+theta)/beta ratio has been found to correlate with a decrease in alertness level (de Waard & Brookhuis, 1991; Sengupta et al., 2017).

Neuroimaging is generally not a field-usable measure, though portable EEG systems are making EEG studies more viable in the field. For example, Ko and colleagues' (2015) study uses a single channel, wireless EEG device for driver fatigue detection in real-time. Because this systematic review aims to inform a cognitive fatigue measure that can be used in a variety of environments and take multiple measurements on a daily basis, it is imperative that individuals are able and confident in completing cognitive fatigue measurements independently, which would be difficult without training, or trained personnel, to carry out the cognitive fatigue readings.

It is important to note that the studies included which have neuroimaging have at least one other objective measure of cognitive fatigue that is not neuroimaging, as per our inclusion criteria. Therefore, neuroimaging results have been discussed alongside the alternative measure of objective cognitive fatigue in the relevant sections.

Physiological Measures

This section discusses the use of physiological measures such as heart rate, skin conductance and facial features for the objective measurement of cognitive fatigue.

ECG & Heart Rate Metrics

Heart rate is an easily accessible metric that is recorded using the number of contractions per minute to determine the speed of an individual's heartbeat. Heart rate is readily affected by circumstances such as sleep, anxiety, physical exercise and stress (Abdullah & Mustapar, 2019). Similarly, changes in the heart rate over a given period of time, known as heart rate variability, allow researchers to study how evolving circumstances affects the heart rate. As previously discussed in the driving measures section, heart rate seems to have too many confounding factors during real-world driving and was only used as a metric in three out of fourteen driving simulation studies, of which only two out of three reported a significant increase.

Nine other studies that used heart rate metrics for measurement of cognitive fatigue included (Ferreira et al., 2020; Gergelyfi et al., 2015; Herlambang et al., 2020; Smith et al., 2019; Sun et al., 2022; Van Cutsem et al., 2019; Wilson et al., 2020; Zhang et al., 2008; Zhang & Yu, 2010). Ferreira and colleagues (2020) used a Stroop task to induce cognitive fatigue. They measured heart rate variability at rest, during the final five minutes of the experiment (25 – 30 minutes) and at four 15-minute intervals following the end of the experiment. For the cognitively fatigued participant group, a significant decrease was found between the baseline resting heart rate and the periods of heart rate monitoring post-experimental condition every 15 minutes. However, no significant difference was found between the baseline and final five minutes of the experimental condition. Equally, the significant decrease found in heart rate at the four 15-minute interval periods were not significantly different from each other. Therefore, these results need to be carefully considered in terms of adding to an evidence base for heart rate and cognitive fatigue tracking. Van Cutsem and colleagues (2019) also utilised the Stroop task for cognitive fatigue induction. Contrary to Ferreira and colleagues' (2020) study, they found a significant increase in heart rate during the Stroop task ($p = 0.008$).

Comparatively, Gergelyfi and colleagues' (2015) results revealed a significant decrease in heart rate between all five sudoku task blocks in comparison to block 1 ($p = 0.0001$), which the authors presume reflects an increase in cognitive fatigue. Herlambang and colleagues (2020) also utilised sudoku as a fatigue inducing task. They used the mid-frequency band of heart rate variability (0.07 - 0.14 Hz) and combined this with measuring the heart power output within this mid-frequency band. These measures have previously been supported as a metric of mental effort by separate literature, with higher power within the mid-frequency band indicates less effort investment by participants, and vice versa (Mulder & Mulder, 1981; Tsao et al., 2013). Time-on-task (sudoku blocks) revealed a significant effect on mid-frequency band power, suggesting that participants became less invested in their cognitive effort levels the longer they were engaged in the task for. Furthermore, the interaction effects between time-on-task and condition (high vs low level intrinsic motivation conditions of the sudoku blocks as described in the cognitive fatigue inducing section) also revealed significant results. Low level intrinsic motivation blocks had a significantly higher mid-frequency band power in comparison to the high-level intrinsic motivation blocks, indicating that participants invested less cognitive effort during the low motivation blocks. Both Gergelyfi and colleagues (2015) and Herlambang and colleagues (2020) used sudoku as their cognitive fatigue inducing task, and both resulted in robust heart rate data for suggesting an increase in cognitive fatigue.

Heart rate also decreased significantly for Zhang and colleagues (2008) who used a switch task to induce cognitive fatigue ($p < 0.005$). Heart power output also significantly increased as time-on-task increased ($p < 0.005$). A similar, but more recent, study undertaken by Zhang and Yu (2010) investigated the same parameters following cognitive fatigue induction via a long-term mental arithmetic task (see *Maths Tasks* under *Cognitive Fatigue Induction*). Again, heart rate was found to decrease significantly between pre- and post-task measurements ($p < 0.005$), with heart power output increasing significantly with time-on-task ($p < 0.005$). Both studies provide results that are in line with Herlambang and colleagues' (2020) study, with increasing power suggesting a decrease in cognitive effort investment from participants due to cognitive fatigue and/or lack of motivation.

Qin and colleagues (2021) analysed heart rate variability during their aviation task to determine changes in objective cognitive fatigue. The ratio of high-frequency (0.15 – 0.4 Hz) and low-frequency (0.04 – 0.15 Hz) bands of power were used as a metric of cognitive fatigue. Pairwise comparisons revealed a significant difference between all six time periods of flight simulation ($p < 0.05$) indicating an increase in cognitive fatigue, and a significant difference was found for the main effect of time-on-task for low- to high-frequency ratio ($p < 0.001$). Wilson and colleagues (2020) also measured heart rate during their aviation tasks, and found a significant decline over the four time points for the mean heart rate between the first (baseline) and last measurement ($p < 0.05$). A decrease was also found between time points one, two and three, however these were not statistically significant.

Sun and colleagues (2022) found that heart rate rose significantly between baseline and post-task ($p < 0.001$) following the use of a continuous performance task (as described in the AX-CPT section in inducing cognitive fatigue). This contradicts the previously discussed studies and suggests that heart rate increase, as opposed to decrease, is indicative of exacerbated cognitive fatigue. Smith and colleagues (2019) also used the AX-continuous performance task, Stroop task or a psychomotor vigilance task to induce cognitive fatigue. Heart rate was measured pre-task, and every ten minutes during the task, as well as a three-minute sample that was taken immediately post-task. No significant results were revealed for heart rate variability for the main effects of task type or for an interaction of task type x time-on-task. The heart power output revealed some significant results; however, they were inconsistent and task dependent.

In summary, while some studies support the hypothesis that heart rate decreases with cognitive fatigue (Ferreira et al., 2020; Gergelyfi et al., 2015; Wilson et al., 2020; Zhang et al., 2008; Zhang & Yu., 2010), others suggest an increase (Sun et al., 2022; Van Cutsem et al., 2019), or show mixed results (Smith et al., 2019). These variations highlight the complexity of using heart rate metrics as a measure of cognitive fatigue and suggest that multiple factors, including task type, individual differences, environmental factors and measurement timing, can influence outcomes.

Three studies utilised ECG following a Stroop, pseudoword task and n-back task (Habay et al., 2021; Herlambang et al., 2019; Tanaka et al., 2011) respectively. Habay and colleagues (2021) found a significant decrease in heart rate from pre- to post-Stroop ($p < 0.001$), but reported no other measures. Herlambang and colleagues' (2019) study revealed a significant increase in mid-frequency band power over time ($p < 0.001$), however, high frequency power did not show significant change. As an increase in mid-frequency power indicates less cognitive effort (Herlambang et al., 2019), it is worth considering that the pseudoword task may have revealed a practice effect through ECG measures, as it would be expected for cognitive effort to increase over time-on-task. Tanaka and colleagues (2011) utilised the n-back task, which revealed a significant decrease in the high-frequency band from pre- to post-task ($p = 0.020$), however the low-frequency band and low to high frequency ratio did not show a significant change.

Jagannath & Balasubramanian (2014) used various physiological measures, including heart rate, surface electromyography, seat interface pressure, blood pressure, EEG and oxygen saturation levels to monitor fatigue. Heart rate was found to decrease significantly with time-on-task, as mentioned previously in the *Driving and Driving Performance* section. Surface electromyography was utilised for evaluation of physical fatigue only and will therefore not be discussed. EEG showed significant differences in expected brain activity for cognitive fatigue between pre- and post-driving task ($p < 0.05$). Significant differences ($p < 0.05$) were also observed for both systolic and diastolic blood pressure pre- and post-driving, however, the authors note that their findings were contradictory to a previous study (Kobayashi et al., 2002) that found an increase in blood pressure over time-on-task. However, the previous study used real-world driving, in comparison to a driving simulator that was utilised by Jagannath & Balasubramanian (2014), and, therefore, could be indicative of confounding factors such as stress levels instead of cognitive fatigue. Blood oxygen saturation levels remained relatively consistent throughout the driving sessions, indicating that this is not a sensitive measure to cognitive fatigue induction following driving simulation. The authors expressed an expectation of oxygen saturation levels to decrease following cognitive fatigue onset due to prior literature searches. Lastly, seat interface pressure was analysed due to correlations found between physical discomfort and fatigue arising from oxygen deficiency due to being in a static position for a prolonged period of time. This study found a significant increase in seat interface pressure due to shifting body movements that resulted in smaller surface area contact between the individual and the seat, which causes higher pressure in specific seat areas. This could be used as an objective measure of fatigue induction, however, being able to distinguish between cognitive and physical fatigue seems too challenging to be able to utilise this metric solely for cognitive fatigue monitoring.

Skin Conductance

Skin conductance allows cognitive fatigue to be monitored through changes in the electrical properties of the skin. When an individual is cognitive fatigued, skin resistance decreases, allowing the electrical skin signal to increase, resulting in an overall increase in skin conductance. Sun and colleagues (2022) found a significant increase in skin conductance between pre- and post- cognitive fatigue induction via a continuous performance task, indicating an increase in cognitive fatigue levels.

Gergelyfi and colleagues (2015) also utilised skin conductance as a cognitive fatigue metric. Due to technical difficulties, eight out of 18 participants were discounted as the data was not correctly recorded. However, the results were analysed with this discrepancy in mind. Higher motivation blocks of the sudoku task due to reward revealed a significant increase (p = 0.05) of skin conductance, indicating a higher investment of cognitive resources to overcome cognitive fatigue due to reward incentive. A significant positive relation between reward effect on accuracy and behaviour was also found when analysing skin conductance, reward, and accuracy when using regression lines due to the smaller sample size for skin conductance (p = 0.021), suggesting that skin conductance increase could be a correlate of motivation.

Plethysmography

Plethysmography is used to measure the absorption of light by haemoglobin, which correlates with blood flow volume. Saito and colleagues (2018), who used the Stroop task for cognitive fatigue induction, did not find any significant differences.

Wilson and colleagues (2020) used plethysmography during their aviation task to monitor participants blood flow. They discuss monitoring heart rate variability features using plethysmography and electrocardiogram, and determined a high correlation between both data sets. The authors suggest that plethysmography could therefore be used as a substitute for electrocardiogram use. The heart rate measures gained from the plethysmography were found to have a significant decrease between baseline and the final time point ($p = 0.05$).

Pattyn and colleagues (2008) used a different type of plethysmography called respiratory inductive plethysmography. This technique is used to measure respiratory movements, in this case, abdominal and ribcage excursions were recorded in order to track tidal volume, respiratory frequency and changes in the R-R heart rate interval. The only significant result found in this study was that of respiratory frequency, which increased significantly from baseline to the first measurement interval ($p = 0.009$), however continued to stay plateaued for the final two measurements over the total 30-minute period of completing a cognitive task. Tidal volume showed the same trend without significant effect, whilst R-R interval showed a non-significant but steady increase as time-on-task increased.

Wilson et al. (2020) found that plethysmography effectively tracked heart rate changes and showed a significant decrease from baseline, indicating its potential as an alternative to ECG. In contrast, Saito et al. (2018) found no significant results with plethysmography for cognitive fatigue, and Pattyn et al. (2008) reported only a significant increase in respiratory frequency, with no substantial changes in tidal volume or R-R interval over time. Therefore, these studies do not indicate that plethysmography is a reliable enough measure for objective cognitive fatigue.

Facial Features

Mehmood and colleagues (2022) additionally utilised facial feature recording as a means to measure objective cognitive fatigue. A colour video camera was mounted within the excavation equipment to record the operators' facial features within the cabin. The facial features analysed were eye aspect ratio and eye distance, eyebrow position, mouth aspect ratio, nose to jaw and nose to chin ratios and face area and head motion. Subjects revealed a significant decrease in eye aspect ratio ($p < 0.05$) over time, with eye distance showing a non-significant increase over time. A significant main effect of time on eyebrow features and mouth aspect ratio was also revealed ($p < 0.05$). A significantly decreasing nose to chin ratio was revealed ($p < 0.05$), however, no significant effects were observed for nose to jaw ratio. Lastly, head motion and facial feature measurement showed significant increases over time ($p < 0.05$ for both). Mehmood and colleagues' (2022) article briefly discusses how eye behaviour can be used to evaluate mental state, due to the eyes being an origin for high information intake to the brain (the authors did not clarify specifically what the individual facial features referred to). It is likely that, in a real-world context, the use of facial features would be too complex to implement due to individual differences in facial responses, difficulties in maintaining baselines, and potential differing interpretations of facial feature movements within literature.

OBJECTIVE TASK PERFORMANCE

It is worth noting for this section that many papers included in this review record task performance data for the task which they are using to actively induce cognitive fatigue. These results will not be elaborated upon because this data essentially reflects the fatiguing effects of the task itself. Using such information to inform the development of an objective cognitive fatigue measure may therefore be unreliable, as the measurement task is actively employed to induce cognitive fatigue. This attribute is not a desirable characteristic in the context of developing a new assessment tool.

Advanced Trail Making Task

One paper used the advanced trail making task to measure objective cognitive fatigue, (Tanaka et al., 2011) as a result of the n-back task. The advanced trail making task was performed for 30 minutes. It required participants to complete a computerised 'dot-to-dot', whereby the mouse is used to select the dots in order from one to 25. Tanaka and colleagues' (2010) had three different conditions for their trail making task. In condition A, the colour of the dot changed from black to yellow subsequent to being clicked on. In condition B, selection of the first target dot caused it to disappear and be replaced with a 26th target dot, as with the second target dot that was replaced with a 27th target dot, and so on. Condition C was the same as condition B, however the position of the new targets was randomised, as well as all remaining targets changing position. All conditions ended when the participant selected the 25th dot in the sequence. Participants completed these conditions A, then B, then C, sequentially for the 30 minutes. The total error count for the advanced trail making task had a tendency to increase after the n-back, however did not reach statistical significance ($p = 0.076$). In comparison, the total trial counts of the advanced trail making task increased significantly ($p = 0.001$), indicating that participants were faster to complete the trail making task trials following the n-back task. These results suggest that either the n-back did not significantly cognitively fatigue participants, and/or that the performance measures for the advanced trail making task were not sensitive enough to capture increased cognitive fatigue. The fact that participants' speed improved for the trail making task following the n-back seems to be indicative of practice effects, or may indicate that the n-back results in modality specific cognitive fatigue. The advanced trail making task is a visual task that does not require a heavy strain on working memory resources, whereas the n-back is a working memory task.

Secondary Tasks

Qin and colleagues (2021) used a secondary 'attacking' task during their study, which utilised flight simulation as a cognitive fatigue inducing task (see *Aviation*). Participants were required to press two buttons on the joystick successively following the appearance of an enemy plane appearing on the display screen. They were not required to aim at the target. However, due to limitations of this task discussed by the authors, the task performance was not analysed.

The impact of cognitive fatigue was assessed using a working memory task in Gergelyfi and colleagues' (2015) study. This working memory task had varying levels of extrinsic motivation through three reward levels (1, 10 and 50 points). To heighten task difficulty, two conditions were randomly selected within each block of 14 trials. The first condition had seven randomly numbers presented between 1 and 8, and participants were required to report as quickly as possible the missing number. The second condition was the same principle, but eight numbers were presented between 1 and 9. No feedback was given on subjects' performance. Data was recorded in terms of 'decision time', operationalised as the length of time taken between the numbers being presented and subjects selecting a missing numeral. A significant main effect of reward was observed for working memory task accuracy ($p < 0.001$), with participants demonstrating higher accuracy with increased incentive. Furthermore, reward also influenced decision time, with participants taking longer to respond to the task when reward incentive was highest ($p < 0.0001$), presumably to make sure their answer was correct.

The digit symbol coding task and the symbol search task were both used in Jonasson and colleagues' (2018) study. Both these tasks are subtests within the WAIS-IV to assess processing speed index (Abdelhamid et al., 2021). Each trial (row) of the symbol search contained two target symbols and five search symbols. 15 rows of trials were presented simultaneously. Participants were instructed to indicate whether either of the target symbols matched any of the five search symbols, with responses (either 'yes' or 'no') made by clicking on a checkbox. Numbers of target-present and target-absent responses were approximately equal. The only significant result reported for the symbol search task was that the control group performed significantly better than the cognitively fatigued group ($p = 0.008$).

In the digit symbol coding task, rows of small blank squares were presented, each accompanied by a randomly assigned number ranging from one to nine. Above these rows, a key is presented that matches each number to an assigned geometric symbol. The subjects were given two minutes to match specific numbers with the corresponding symbols using the provided key. Performance is assessed with a scaled score of correctly matched numbers and symbols. Significant interaction effects were found for the digit symbol coding task when comparing pre- and post-test scores and group differences ($p = 0.003$).

Claros-Salinas and colleagues (2010) used a simple reaction time task and a divided attention task as further secondary tasks to the Go/NoGo task (*see Go/NoGo* under *Inducing and Measuring Cognitive Fatigue*). The reaction time task had a total of 40 trials and lasted approximately three minutes. A white cross was displayed following a random inter-trial interval and participants were required to press a response key as quickly as possible following the stimuli appearance. The divided attention task lasted approximately four minutes and employed a visual and auditory dual-task paradigm. The visual aspect of the paradigm had white crosses appear in a random orientation within a 4x4 matrix, changing every one second. In 17 out of the total 100 configurations, the crosses formed the target configuration, a 2x2 square, which participants were instructed to detect and press a response key. Simultaneously, the auditory aspect of the paradigm was completed. 200 tones were presented successively with two different frequencies for 433ms. While the frequency typically changed between subsequent tones, 16 tone sequences were kept the same. Participants were required to press the response key upon detecting a repetition in the tone frequency. Performance measures included reaction times and errors (both missed responses and incorrect responses). Task performance for both the auditory and visual aspects of this paradigm were largely consistent. Control participants demonstrated significantly better task performance in comparison to stroke and multiple sclerosis patients ($p < 0.001$). Performance for both the visual and auditory aspects also showed improvement over the two experimental days for the control group, in comparison to the patient group where performance remained consistent. Similarly, error rates were significantly reduced in the control vs patient groups ($p < 0.01$).

Gergelyfi and colleagues (2015) found that increased extrinsic motivation improved accuracy on a working memory task but also led to longer decision times, suggesting a trade-off between speed and accuracy with higher incentives. Claros-Salinas and colleagues (2010) observed that control participants outperformed stroke and multiple sclerosis patients in both reaction time and divided attention tasks, suggesting that the patient group struggled more with speed and accuracy over time. Jonasson and colleagues (2018) reported that a cognitively fatigued group performed worse on the symbol search task and showed significant interaction effects in the digit symbol coding task, indicating cognitive fatigue impacts processing speed and accuracy. These studies cumulatively help evidence that speed and accuracy can be used as markers for objective cognitive fatigue through task performance of secondary tasks.

Maths Tasks

Four studies over three articles utilised a maths task (Chen et al., 2021; Hu & Lodewijks, 2021; Liu & Wu, 2008). Both Hu & Lodewijks' (2021) experiments used a number calculation task. Participants listened to an audio task which played a random number between 1 and 9 every three seconds whilst driving. The participants were required to complete addition of two consecutive numbers, the number presented and the number that occurred immediately prior. They then spoke the answer aloud. These trials ran for five minutes for every ten-minute driving time-on-task interval. This secondary task was completed to add to the participants' workload, however, data was not collected. Therefore, this task was not used to objectively measure cognitive fatigue.

Liu and Wu (2008) employed an arithmetic task by having participants answer two addition and two subtraction questions orally whilst driving for cognitive fatigue induction. No significant differences were found in performance accuracy, however, different road conditions had an effect on response times. Monotonous road conditions resulted in slower response times than complex road conditions ($p = 0.020$).

In Chen and colleagues' (2021) study, the interruption task condition included participants completing a secondary maths task whereby addition within 100 was required. The addition was displayed on screen for 500ms, followed by a fixed cross for 2000ms. This was repeated five times in total during the interruption trial phase. Accuracy and reaction times were significantly reduced in the fatigue condition in comparison to the non-fatigue condition, indicating that an interruption to the primary task required more cognitive resources in order to redirect attention back to the primary task.

Vigilance Tasks

Seven studies used a vigilance task to measure objective cognitive fatigue induction (Feyer et al., 1994; Hu & Lodewijks, 2021; Li et al., 2004; Naeeri et al., 2021; Qin et al., 2021; Smith et al., 2019; Williamson et al., 1996). Both Hu and Lodewijks' (2021) studies used a screen display task for five minutes pre- and post- one of three driving conditions (see *Table 6*) to track vigilance by utilising eye tracking software. A white dot was used as a fixation target on a grey square, that was displayed on a computer screen in front of the participants. The white dot stimuli randomly emerged from a corner of the grey square. A subsequent saccadic target would also then be presented randomly at one of seven distances from the white dot, and as one of three random directions (either a horizontal, vertical, or diagonal line). Participants were required to press 'K' on a keyboard when the white dot stimuli appeared on the edge of the grey square. The authors found only a significant difference in their one-hour overload driving condition; the one-hour underload and the 1.5-hour overload driving conditions did not yield significant differences between pre- and post-task measurements.

Li and colleagues (2004) used a vigilance task when exploring the use of magnitopuncture during driving conditions (see *Critical Flicker Fusion Task*). Magnitopuncture is similar to acupuncture, however, instead of using needles as in traditional acupuncture practice, the acupuncture points are manipulated via magnetic fields. Participants were presented with three random numbers, displayed simultaneously. The combination of numbers changed every second. Subjects were required to click the right mouse button upon recognition of three different, odd numbers being displayed simultaneously. The control group, who were not subjected to fatigue relieving magnitopuncture, had significantly increased reaction times ($p < 0.01$) in the pre- and post-task measurement, suggesting a decrease in vigilance as a result of increased cognitive fatigue.

Feyer and colleagues (1994) also used a vigilance task to assess the impact of driving on cognitive fatigue. Their task setup included a semi-circular display with five lights, each accompanied by a corresponding button directly below. An extra light and button were positioned equidistantly between each endpoint and the top of the semi-circle. The lights lit up individually in a quasi-random sequence, with the participants' goal to hover over the display and press the button corresponding to the illuminated light. Occasionally at a predetermined frequency, two lights were lit simultaneously, requiring the subject to press the central bottom button. The number of errors made and the speed of accurate responses were recorded. There was a significant main effect of time-on-task on response speed ($p = 0.001$) indicating that vigilance performance consistently improved towards the end of the trip, with no significant effects found for percentage of errors made.

During Qin and colleagues' (2021) study, aviation was used to induce cognitive fatigue (see *Aviation*), during which a secondary attacking task took place (see *Secondary Tasks*). Once an 'attack' had commenced, a simultaneous flashing signal light and warning voice would occur, which participants needed to respond to by pressing a button on the joystick. Reaction times, missed trials and false starts (pre-emptive responses) were recorded. Reaction times showed a significant increase over time-on-task ($p = 0.013$), suggesting a decrease in vigilance as time-on-task increased. False alarms and missed trials were not statistically analysed due to the low occurrence rate.

Williamson and colleagues (1996) used a vigilance task pre- and post-driving task by having participants press the necessary button to extinguish a light stimulus. Over a 10-minute period, 600 light stimuli were presented sequentially. No further information was provided on the specifics of this vigilance test. The reaction time and number of errors were recorded.

Significant decreases were found between pre- and post-driving reaction times, suggesting an increase in alertness ($p < 0.002$). The mean frequency of errors did not change.

The psychomotor vigilance task was used in two of the seven studies (Naeeri et al., 2021; Smith et al., 2019). It is a simple visual reaction time task whereby a timer is presented in the middle of the screen, which counts in millisecond intervals. The timer continues until the participant presses the response button, resulting in their reaction time, or 500 ms had passed, at which point a "missed" message is displayed. In order to rectify pre-emptive or repetitive response button clicking, reaction times less than 100 ms were met with a 'false start' error message. Smith and colleagues (2019) used this vigilance task for three minutes pre- and post- one of three cognitively fatiguing tasks that lasted for 45 minutes (Stroop, AX-continuous performance task, and the psychomotor vigilance task). They found significant main effects of time ($p = 0.003$) and task type ($p = 0.024$) between the baseline and post-psychomotor vigilance task performance. No significant difference was found for the interaction effect of time x task type.

Naeeri and colleagues (2021) also used the psychomotor vigilance task following their aviation task for inducing cognitive fatigue, whereby expert and novice pilots were recruited (see *Aviation*). The authors found that mean reaction times were significantly faster, with lower missed trials and false starts when comparing expert and novice pilots ($p < 0.001$ for all three statistics). This suggests the novice pilots were subjected to higher vigilance decrement due to cognitive resource allocation to the aviation task being higher, and as a result, are likely to have succumbed to a higher level of cognitive fatigue.

Overall, the studies using vigilance tasks to measure cognitive fatigue yielded varied results. Hu & Lodewijks (2021) found significant vigilance effects only under specific driving conditions. Li and colleagues (2004) observed increased reaction times in the control group, suggesting decreased vigilance without magnitopuncture. Qin and colleagues (2021) noted increased reaction times over time-on-task during aviation tasks, while Williamson and colleagues (1996) found improved alertness post-driving with decreased reaction times but unchanged error rates. Both Smith and colleagues (2019) and Naeeri and colleagues (2021) highlighted significant effects of cognitive fatigue on vigilance. The variation in the results is likely to be due to the variation in the vigilance tasks themselves, as well as the difference in the primary tasks being utilised to induce cognitive fatigue.

Visuomotor and Psychomotor Performance

Two studies utilised visuomotor tasks, with Habay and colleagues' (2021) study utilising a visuomotor task based on Van Cutsem and colleagues' (2019) model. The setup of this visuomotor task was as follows: eight LED lights, six on a table, one behind the participant's preferred hand, were used. The last light was kept in a dark box. Participants extinguished lit lights based on colour cues, either green, red or yellow with a table tennis paddle (simple stimuli). When the light turned blue, they were instructed to ignore the stimuli in front and extinguish the light situated on their preferred hand side (complex stimuli). Colours appeared randomly 16 times each, totalling 64 stimuli. Inter-stimulus time varied between three and six seconds for a 6-minute duration. Accuracy and response time were recorded for each stimulus type. Participants missing over 25% of stimuli in this visuomotor task were excluded from further analysis. A significant interaction effect was found between time and condition (cognitive fatigue induced via the Stroop task vs documentary watching for the control group; $p = 0.050$). Additional post-hoc testing revealed a significantly increased reaction time from pre- to post-task in the cognitively fatigued condition for all stimuli ($p = 0.006$), with a significant difference found between the post-task reaction times of the inhibitory (complex) stimuli between the cognitively fatigued and control groups ($p = 0.001$).

Van Cutsem and colleagues' (2019) study used seven lights, set up against a wall, that lit up either red, blue green or yellow. If a light turned red, green or yellow, participants were instructed to wave their hand in front of the light within a 5cm range to extinguish it (simple stimuli). However, if a light turned blue, they were instructed to extinguish a light positioned behind them as fast as possible (complex stimuli). As with Habay and colleagues (2021), each colour appeared randomly 16 times each, totalling 64 stimuli. Inter-stimulus time ranged between three and six seconds and were used an equal number of times, with total task duration lasting approximately six and a half minutes. Accuracy and response time were recorded for each stimulus type. No effect of time or condition (cognitive fatigue induced via the Stroop task vs documentary watching for the control group) was found for accuracy, however, an interaction effect of condition x time-on-task was found for response time ($p = 0.011$).

Participants reaction times became significantly increased for the cognitively fatigued group only ($p < 0.001$).

Habay and colleagues (2021) and Van Cutsem and colleagues (2019) both found that cognitive fatigue induced by the Stroop task significantly increased reaction times in visuomotor tasks, particularly for complex stimuli. However, neither study observed an effect of cognitive fatigue on task accuracy. This indicates a speed-accuracy trade-off that could correlate to a decline in objective cognitive fatigue.

Conversely, Völker and colleagues (2016) used a psychomotor task that required participants to touch a red dot as quickly as possible with their dominant hand. The dot was displayed on a black background via mobile phone. Touching the dot removed it from the display and reaction times were recorded. The next trial began 0-8 seconds following prior trial completion. Outcome measures included the mean and standard deviation of reaction times, as well as the count of premature responses (i.e., touching the display before the red dot appeared). The task lasted for a total of 300 seconds. No significant results were found for this task for reaction time mean or standard deviation, and the authors concluded that it was therefore not a valid indicator of cognitive fatigue.

Eye Tracking and Pupillometry

Research into eye-tracking and pupillometry has gained traction due to the advancements in equipment, allowing it to become more readily available to researchers. Moreover, eye-tracking and pupillometry is a non-invasive measure that can be used continuously throughout a task to accumulate data without causing exacerbation of cognitive fatigue or requiring a break in the task (Bafna & Hansen, 2021). Table 7 below shows an overview of pupillometry and eye-tracking metrics for all studies included within this review.

Table 7

An overview of all studies that utilised pupillometry and eye-tracking metrics, including result significance

Metric	*Study*	Time on Task (min)	Cognitive Fatigue induction Task	Increase or decrease in relation to cognitive fatigue	Significant?	Comments
Pupil Size/Diameter						
	LeDuc et al. (2005)	72	Helicopter Flight	Increase	Yes	
	Gergelyfi et al., (2015)	120	Sudoku	N/A	N/A	
	Hopstaken et al. (2015)	60	n-back	Decrease	Yes	
	Hopstaken et al. (2015)	60	n-back	Decrease	Yes	
	Körber et al. (2015)	42.5	Driving	Decrease	Yes	
	Hopstaken et al. (2016)	90	n-back	Decrease	Yes	
	Ma et al. (2018)	60	Simulated Driving	Decrease	Yes	
	Herlambang et al. (2019)	150	Vowel Counting Pseudoword Task	Decrease	Yes	Significant decrease for time-on-task; significant increase for reward blocks
	Herlambang et al. (2020)	39	Sudoku	Decrease	Yes	
	Hu & Lodewijks (2021a)	60	Simulated Driving	Decrease	Yes	
	Hu & Lodewijks (2021b)	60	Simulated Driving	Decreases	Yes	Significant for 1 hour underload driving condition; Not significant for 1.5 hour overload driving condition
	Naeeri et al. (2021)	60	Flight Simulator	Decrease	Yes	
	Qin et al. (2021)	90	Simulated flight	N/A	N/A	
	Guillemin et al. (2022)	32	TloadDback	Increase	Yes	"only a condition effect was observed, with larger pupil size in the HCL condition"
	Ioannucci (2022a)	41	Auditory task (either: n-back task, a pitch-sequence reproduction task, or a switch task)	Decrease	Yes	Decrease for all auditory task types
	Ioannucci (2022b)	41	Auditory task (either: n-back task, a pitch-sequence reproduction task, or a switch task)	Increase	Yes	"larger pupil size corresponded to greater performance drop in TDT"; auditory task relationship not discussed
	Ioannucci (2022c)	41	Auditory task (either: n-back task, a pitch-sequence reproduction task, or a switch task)	Increase	Yes	"larger pupil size corresponded to greater performance drop in TDT"; auditory task relationship not discussed
	Shi et al. (2022)	90	Simulated driving	Decrease	Yes	Significant decrease for minimum pupil diameter for TOT; Not significant results for maximum pupil diameter for TOT

Metric	*Study*	Time on Task (min)	Cognitive Fatigue induction Task	Increase or decrease in relation to cognitive fatigue	Significant?	Comments
	Sun et al. (2022)	120	Continuous Performance Test	N/A	N/A	
	Yu et al. (2022)	180	Go/NoGo	Increase	Yes	For Go trials only
	Zhao et al. (2022)	60	Driving simulator	Decrease	Yes	
Pupil Dilation						
	Hopstaken et al. (2015)	60	n-back	Decrease	Yes	
	Zargari et al. (2018)	40	Pattern Recognition Task	Increase	Yes	
	Qin et al. (2021)	90	Simulated Flight	N/A	N/A	
Pupil Dilation Velocity						
	Shi et al. (2022)	90	Simulated Driving	N/A	N/A	
Pupil Constriction Latency						
	LeDuc et al. (2005)	72	Helicopter Flight	Increase	Yes	
	Shi et al. (2022)	90	Simulated Driving	N/A	N/A	
Pupil Constriction Amplitude						
	LeDuc et al. (2005)	72	Helicopter Flight	N/A	N/A	
Pupil Constriction Velocity						
	Shi et al. (2022)	90	Simulated Driving	Increase	Yes	
Eye Blink Frequency						
	Gergelyfi et al., (2015)	120	Sudoku	Increase	Yes	
	Körber et al. (2015)	42.5	Driving	Increase	Yes	
	Völker et al. (2016)	480	Working Day (Clerical)	N/A	N/A	
	Zhao et al. (2017)	90	Simulated Driving	Increase	Yes	
	Zargari et al. (2018)	40	Pattern Recognition Task	Increase	Yes	
	Herlambang et al. (2019)	150	Vowel Counting Pseudoword Task	Increase	Yes	Significant increase for time-on-task; significant decrease for reward blocks
	Herlambang et al. (2020)	39	Sudoku	Increase	Yes	
	Hu & Lodewijks (2021a)	60	Simulated Driving	N/A	N/A	
	Hu & Lodewijks (2021b)	60	Simulated Driving	N/A	N/A	
	Qin et al. (2021)	90	Simulated Flight	Increase	Yes	
	Guillemin et al. (2022)	32	TloadDback	Increase	Yes	Only long blinks counted; "number of long blinks revealed a significant Condition*Group interaction; post-hoc analysis evidenced an effect of condition in both groups with higher numbers of blinks during the high cognitive load condition"

	Lampe & Deml (2022)	40	Driving	N/A	N/A	
Metric	***Study***	**Time on Task (min)**	**Cognitive Fatigue induction Task**	**Increase or decrease in relation to cognitive fatigue**	**Significant?**	**Comments**
	Sun et al. (2022)	120	Continuous Performance Test	N/A	N/A	
Blink Duration						
	Gergelyfi et al., (2015)	120	Sudoku	N/A	N/A	
	Körber et al. (2015)	42.5	Driving	Increase	Yes	
	Zargari et al. (2018)	40	Pattern Recognition Task	Increase	Yes	
	Herlambang et al. (2019)	150	Vowel Counting Pseudoword Task	Increase	Yes	Significant for TOT; No significant effect for reward blocks
	Herlambang et al. (2020)	39	Sudoku	Increase	Yes	
	Hu & Lodewijks (2021a)	60	Simulated Driving	Increase	Yes	Significant for active cognitive fatigue induction only
	Hu & Lodewijks (2021b)	60	Simulated Driving	Increase	Yes	
	Lampe & Deml (2022)	40	Simulated Driving	Increase	Yes	
Saccade Frequency						
	Herlambang et al. (2019)	150	Vowel Counting Pseudoword Task	Increase	Yes	Significant increase for TOT (although graph shows decrease); effect of reward was not significant; Reward*Time had significantly higher saccades frequency in reward than in non-reward blocks.
	Herlambang et al. (2020)	39	Sudoku	Decrease	Yes	Used as an indicator of attention
Saccade Velocity						
	LeDuc et al. (2005)	72	Helicopter Flight	Decrease	Yes	
	Zargari et al. (2018)	40	Pattern Recognition Task	Decrease	Yes	
	Hu & Lodewijks (2021a)	60	Simulated Driving			
	Hu & Lodewijks (2021b)	60	Simulated Driving	N/A	No	
Saccade Duration						
	Zargari et al. (2018)	40	Pattern Recognition Task	Decrease	Yes	
	Hu & Lodewijks (2021a)	60	Simulated Driving	Increase	Yes	
	Hu & Lodewijks (2021b)	60	Simulated Driving	Increase	Yes	
Saccade Amplitude						
	Herlambang et al. (2019)	150	Vowel Counting Pseudoword Task	Increase	Yes	Significant increase for TOT; effect of reward was not significant; Reward*Time had significantly higher saccades amplitude in non-reward than in reward blocks.

Metric	Study	Time on Task (min)	Cognitive Fatigue induction Task	Increase or decrease in relation to cognitive fatigue	Significant?	Comments
	Herlambang et al. (2020)	39	Sudoku	Decrease	Yes	Used as an indicator of attention
% of Eye Closure						
	McKinley et al. (2011)	30	Flying an unmanned aerial vehicle target acquisition task	N/A	N/A	
	Körber et al. (2015)	42.5	Driving	N/A	N/A	
Approximate Entropy for Pupil Position						
	McKinley et al. (2011)	30	Flying an unmanned aerial vehicle target acquisition task		Yes	
	Naeeri et al. (2021)	60	Flight Simulator	Increase	Yes	
Gaze Positions						
	Hopstaken et al. (2016)	90	n-back	Increase	Yes	Significant increase for off-screen and missing gaze position for TOT; significant increase in on target gaze during reward manipulation
Visual Distraction Frequency						
	Herlambang et al. (2019)	150	Vowel Counting Pseudoword Task	Increase	Yes	
	Herlambang et al. (2020)	39	Sudoku	Increase	Yes	
Visual Distraction Duration						
	Herlambang et al. (2019)	150	Vowel Counting Pseudoword Task	Increase	Yes	
	Herlambang et al. (2020)	39	Sudoku	Increase	Yes	
Eye Fixation Number						
	Naeeri et al. (2021)	60	Flight Simulator	Decrease	Yes	
Eye Fixation Duration						
	Zargari et al. (2018)	40	Pattern Recognition Task	Increase	Yes	
	Naeeri et al. (2021)	60	Flight Simulator	Increase	Yes	
% Pupillary Light Reflex						
	Shi et al. (2022)	90	Simulated Driving	Increase	Yes	
Eyelid Distance						
	Lampe & Deml (2022)	40	Simulated Driving	Decrease	Yes	

Various studies recorded phasic changes in pupil diameter and pupil diameter range to determine if it can be used as a measure of objective cognitive fatigue (Gergelyfi et al., 2015b; Guillemin et al., 2022; Herlambang et al., 2019, 2020; Hopstaken et al., 2016; Hopstaken et al., 2015; Hu & Lodewijks, 2021; Ioannucci et al., 2022; Körber et al., 2015; LeDuc et al., 2005; Zargari et al., 2018; Naeeri et al., 2021; Qin et al., 2021; Shi et al., 2022; Sun et al., 2022; Yu et al., 2022; Zhao et al., 2022). The tasks used to induce cognitive fatigue in these studies ranged from 39 minutes to 120 minutes. Of the 20 experiments that measured pupil diameter, one study did not yield significant results (Qin et al., 2021a), and five studies showed a significant increase in pupil diameter. The remaining thirteen had significant decreases in pupil size as time-on-task increased. Hu & Lodewijks (2021) first experiment showed significant effects for one hour of passive driving on pupil diameter, but their second experiment which required participants to drive for 1.5 hours did not produce a significant effect. Yu and colleagues' (2022) study also revealed a significant effect under certain conditions; the Go/NoGo paradigm used resulted in significant changes in pupil diameter during the Go trials only.

For Ioannucci and colleagues (2022), pupillometry was used as a secondary measure to the task performance throughout the experiments confirmed that the difference in conditions had a physiological impact, with pupil sizes increasing significantly from the 'easy' to 'hard' conditions, despite being an auditory task. Pupillometry being used as a measure of effort dates back many years within the literature, (Beatty, 1982) however, it has more recently been linked specifically to the auditory modality as an objective measure of listening effort. It has been found that pupil dilation is a physiological response that can occur alongside stress hormones and galvanic skin response to indicate acoustic challenges for the listener (Peelle, 2018). Therefore, the dilation in participants following the 'hard' condition is a reliable indicator that a higher cognitive demand is required, resulting in cognitive fatigue.

LeDuc and colleagues' (2005) study and Shi and colleagues' (2022) study both looked at the pupillary light reflex, of which LeDuc and colleagues' (2005) study found that cognitive fatigue resulted in delayed pupillary response and decreased pupil size following a one-hour aviation task to induce cognitive fatigue. These results were reflected in Shi and colleagues' (2022) more recent findings, whereby the minimum pupil size was significantly decreased, and the pupil constriction percentage significantly increased as a result of the pupillary light reflex following a 90-minute driving task to induce cognitive fatigue. LeDuc and colleagues (2005) showed a significant increase in pupil constriction latency, but Shi and colleagues (2022) did not have a significant finding. Shi and colleagues (2022) also did not have a significant finding for pupil dilation velocity but found a significant increase in pupil constriction velocity. Lastly, LeDuc and colleagues (2005) did not find a significant effect on pupil constriction amplitude. LeDuc and colleagues (2005) did not have a measurement for percentage pupillary light reflex, but Shi and colleagues (2022) found a significant increase in this metric.

Eye blink frequency has been shown to increase with cognitive fatigue due to an increase in time-on-task. Several studies included in this systematic review help to evidence this relationship (Gergelyfi et al., 2015; Guillemin et al., 2022; Herlambang et al., 2019, 2020; Körber et al., 2015; Zargari et al., 2018; Qin et al., 2021; Zhao et al., 2017). Similarly, eye blink duration has also been shown to increase significantly with extended time-on-task (Herlambang et al., 2019, 2020; Hu & Lodewijks, 2021; Körber et al., 2015; Lampe & Deml, 2022; Zargari et al., 2018) and has been strongly evidenced within the included papers. Of the nine studies that recorded blink duration, only one study did not find a significant effect of time-on-task (Gergelyfi et al., 2015a).

Percentage of eye closure (whereby the eyes are more than 80% closed) was observed in two studies (Körber et al., 2015; McKinley et al., 2011). Körber and colleagues' (2015) study used driving simulation to induce cognitive fatigue, whereas McKinley and colleagues (2011) required participants to fly an unmanned aerial vehicle. Neither study found significant results, however, time-on-task was fairly limited in both (42.5 minutes and 30 minutes respectively), so further investigation with an extended time-on-task may help to increase the likelihood of finding results due to potentially small effect sizes (neither study reported effect sizes for this metric). Furthermore, recent studies have generally found percentage of eye closure to be more indicative of drowsiness rather than cognitive fatigue (Abe, 2023).

Eye tracking has also been researched thoroughly in relation to cognitive fatigue, with gaze position and saccades being the primary components of objective measure within this. Gaze position can be further categorised into eye fixation number and duration, and visual distraction frequency and duration. A total of five studies (Herlambang et al., 2019, 2020; Hopstaken et al., 2016; Zargari et al., 2018; Naeeri et al., 2021) observed at least one of these gaze metrics (see Table 7), all of which had a significant result from a time-on-task effect. Herlambang and colleagues' (2019; 2020) studies both saw a significant increase in visual distraction frequency and visual distraction duration, and Hopstaken and colleagues (2016) similarly saw an increase in off-screen and missing target gaze positions as time-on-task increased. This suggests a lack of attention, concentration and motivation as a result of cognitive fatigue from extended time-on-task. Naeeri and colleagues (2021) and Zargari and colleagues (2018) found a significant increase in eye fixation duration following cognitive fatigue induction, suggesting the need to focus more cognitive resources on a task in order to respond as required. Finally, Naeeri and colleagues (2021) also found a significant decrease in eye fixation frequency, suggesting the eye movement was less regular within a given time frame due to the longer duration of eye fixation, which could be either on or off target.

Lastly, saccades can be further categorised into saccade frequency, velocity, duration and amplitude. Five studies collected data on these metrics (Herlambang et al., 2019, 2020; Hu & Lodewijks, 2021; LeDuc et al., 2005; Zargari et al., 2018). Both experiments undertaken by Hu & Lodewijks (2021) recorded data on saccade velocity and duration. Significant increases for saccade duration were found in both experiments, however, only active cognitive fatigue induction from a driving simulator (as opposed to passive simulated driving) resulted in a significant increase in experiment one for saccade velocity. The second experiment did not yield a significant result. The two other studies that explored saccade velocity (leDuc et al., 2005; Zargari et al., 2018) found a significant decrease, contradicting both of Hu and Lodewijks (2021) experimental results. Zargari and colleagues (2018) also found a significant decrease in saccade duration, once again contradicting both experimental results from Hu and Lodewijks (2021). Herlambang and colleagues (2019; 2020) explored saccade frequency and amplitude in both studies, however found different relationships for both metrics. Their 2019 study resulted in a significant decrease in saccade frequency and a significant decrease in saccade amplitude, however, their 2021 study observed a significant increase in both saccade frequency and amplitude. As a result, more literature needs to be considered in order to determine if saccade metrics are a sensitive measure to cognitive fatigue due to the number of contradictions throughout the aforementioned data.

CHAPTER THREE: A DISCUSSION ON MEASUREMENT DEVELOPMENT

The aim of this systematic review is to give an insightful overview of objective cognitive fatigue measures that are used throughout literature, to help inform further research for the development of an objective measure of cognitive fatigue that could be used in 'real-life'. By compiling key information from the results, patterns within the data have emerged, allowing a more robust review of objective cognitive fatigue measures. Inclusion criteria have helped to streamline differences between the studies in order to reduce confounding factors, such as a minimum time-on-task for cognitive fatigue induction being equal to, or more than, 30 minutes. Similarly, assessing the general population with a minimum threshold age of 18 years allows reduction of brain development factors throughout childhood and adolescence that may have resulted in major differences in response to cognitive fatigue inducing tasks. However, this does not rule out the confounding factor of age, and studies that compare between younger and elderly participants have shown significant differences between cognitive fatigue levels (Zargari et al., 2018).

In order to develop a task which is suitable for a large proportion of the population, several factors need to be considered. First, the task needs to be reliable for the objective measurement of cognitive fatigue. The results found from the literature will be discussed surrounding this to identify measures that seem most reliable. Secondly, the task needs to be non-fatiguing in itself. A task that exacerbates fatigue is not appropriate, particularly for use by patients who are already cognitively fatigued and/or subjects who are returning to the workplace. It also makes reliable comparisons across groups challenging, as the task could skew data due to individual differences. Thirdly, the measure needs to be non-intrusive. As participants could be completing this task several times a day, it should be time efficient and not require the use of specialist equipment that may require a change in location and/or be time-consuming to implement. It would also be beneficial if a quiet location was not a necessity based on the range of settings that the measure could be utilised in, as a quiet location may not always be a possibility. Lastly, it needs to be accessible. Accessibility needs to be considered from two angles. The task needs to be easily understandable for a variety of individuals who may have resulting neurological conditions from brain injury or stroke, as well as for generalised individual differences across the population, such as age and intellect. Accessibility also needs to be considered in terms of availability: would it require monetary investment from the healthcare sector or individual for equipment or software? Would it require a trained individual to install or deliver the measure/equipment, or further analyse data following data collection? These questions in particular are barriers to making the measure available to a large population and are therefore vital to consider.

Providing that the criteria outlined above has been met to a reasonable degree, we can then consider design implementation and discuss the feasibility of developing the measure in real-world terms. The above criteria will therefore be used hierarchically, as summarised below, to filter out tasks that would not be appropriate, reliable or realistic for the future development of an objective measure of cognitive fatigue.

The task criteria are discussed in the following hierarchical order:

1. Reliability;
2. Fatigue inducing;
3. Non-intrusive;
4. Accessibility;
5. Design implementation.

RELIABILITY

Reliability is arguably the most important criteria that tasks need to meet in order to be considered further for the development of an objective cognitive fatigue measure. If the task is not reliable and sensitive enough to determine between different levels of cognitive fatigue then it is meaningless to pursue it. For this reason, we had the exclusion criteria of reliability and validity of any measure included within the studies by requiring a second objective measure to be present to help reduce the number of studies with less reliable or valid objective measures.

ECG & Heart Rate

The primary physiological metric used in the studies included in this systematic review was heart rate and heart rate variability. However, results were varied across the studies, bringing into question the reliability of the measure for realistic day to day use. Because heart rate is an easily accessible metric, it has high potential for realistic in-field use. However, heart rate seems to have a number of environmental, emotional, physical and psychological confounding factors, such a stress, anxiety and exercise (Abdullah & Mustapar, 2019), which may be the reason for the variation in results across the studies. As a result, any consequent data needs to be carefully considered before attributing it immediately to objective cognitive fatigue alone. For this reason, it is unlikely to be a reliable measure without significant further analysis of obtained data that can follow research guidelines that are far more evidenced and advanced than is currently available.

Plethysmography

Only two papers in this review utilised plethysmography, one of which recorded blood flow and another which recorded changes in respiratory characteristics. Wilson and colleagues (2020) used blood flow plethysmography, and the heart rate measures gained from the plethysmography were found to have a significant decrease between baseline and the final time point. Pattyn and colleagues (2008) explored respiratory plethysmography, and found that respiratory frequency increased significantly from baseline to post-task, but found no other significant results in the other respiratory characteristics measured. The theory behind plethysmography for both applications is interesting, as it is essentially monitoring the increase in cognitive resource needs. This is accomplished either via the increased intake of oxygen or via increased blood flow that provides increased oxygen supplies and results in the ability for the body to uptake greater levels of glucose in a shorter time period (Kim et al., 2013; Scholey et al., 2020). However, as only two studies utilised this measure in two different contexts, further research would be required to determine if plethysmography can be cognitive fatigue specific, without the confounding emotional, physical and physiological factors that can affect physiological measures (Abdullah & Mustapar, 2019).

Eye-Tracking & Pupillometry

The exploration of eye-tracking and pupillometry has gained increased attention due to advancement in equipment, making it more reliable in terms of data collection, as well as more accessible for researchers. The results from this review indicate that, in particular, eye blink frequency and eye blink duration are reliable indicators of cognitive fatigue. Eight out of thirteen studies showed a significant increase in eye blink frequency, and seven out of eight studies revealing a significant increase in blink duration, as time-on-task increased. However, all of the other pupillometry features (see *Table 7*) did not reveal a strong correlation between increasing or decreasing characteristics of the features, with many studies not reaching significance.

Eye-tracking and pupillometry methods can provide a relatively non-intrusive means of continuously collecting data without requiring task breaks or exacerbating cognitive fatigue in a large number of settings, such as desk work or driving (Bafna & Hansen, 2021). However, for occupations that require constant movement, such as healthcare professionals or construction workers, interruptions of the work day would be required for data collection. Similarly, in many workplace settings, the consistent use of a video camera may infringe on the privacy of surrounding individuals, and active participants may also feel uneasy regarding this method of data collection. Moreover, poor lighting and wearing glasses may result in further unreliability of measures. Additionally, the monetary investment needed to obtain and install the equipment and software, as well as the personnel investment needed to help with this process and analyse the collected data, does not make eye-tracking a feasible option, regardless of the reliability of certain features.

Facial Features

Mehmood and colleagues (2022) utilised facial feature recording as a means to measure objective cognitive fatigue. Subjects revealed a significant decrease in eye aspect ratio, a significant effect on eyebrow features and mouth aspect ratio, a significantly decreasing nose to chin ratio, and significant increases in head motion and facial feature measurement over time. However, only this study utilised facial features within this review, and therefore an overview of the reliability, despite promising results from Mehmood and colleagues (2022), cannot be adequately discussed here. Furthermore, facial feature measurement incurs many of the same issues discussed for eye-tracking and pupillometry, regarding practicalities, privacy and cost due to equipment and personnel requirements.

Skin Conductance

Skin conductance also provides the general advantage of being a non-intrusive physiological measure, meaning that data can be obtained passively throughout the day. However, only two studies utilised skin conductance within this review, and varying conclusions were made. Sun and colleagues (2022) observed a significant rise in skin conductance from pre- and post-cognitive fatigue induction during a continuous performance task, suggesting a relationship between skin conductance and increase in cognitive fatigue levels. However, Gergelyfi and colleagues (2015) additionally manipulated motivation whilst recording skin conductance during the fatiguing task (sudoku). Higher motivation trials revealed a significant increase in skin conductance. The subsequent analysis of skin conductance, reward and accuracy also revealed a significant positive relationship between the reward effect on accuracy and behaviour, strongly suggesting that the increase in skin conductance could be linked to motivation levels. As skin conductance is considered to be an indirect measure of sympathetic autonomic activity, it has been closely correlated to factors such as emotion and attention (Gatti et al., 2018; Laine et al., 2009). In this case, the link to attention is particularly important, as this indicates a change in cognitive resource allocation. Motivation can result in greater attention due to the allocation of cognitive resources, meaning that task engagement can change as a result of subjective goal-directed behaviour (Zhang et al., 2012). Therefore, Gergelyfi and colleagues' (2015) study indicates that the objective physiological measure of skin conductance can be impacted by subjective feelings of cognitive fatigue.

Although various physiological measures show some level of reliability, and certainly have the potential to become more reliable as research develops, many face significant issues surrounding practical implementation. Furthermore, physiological measures may be misleading in terms of comparisons between patient groups and the wider population: many patients will have physiological measures that do not sit within 'normal range' and may not realistically be comparable with others of the same patient group due to individual differences. For these reasons, physiological measures are not recommended for use in the further development of an objective cognitive fatigue measure at this time.

Modality Specific Fatigue - Visual

Reliability of eye-tracking or pupillometry metrics (*see Table 7*) could be compromised due to being a consequent factor of visual fatigue rather than overall cognitive fatigue (Souchet et al., 2022). Many studies reported a decrease in correct responses and reaction times of visual tasks used to measure cognitive fatigue, such as the texture discrimination task and the Flanker task. However, the decrease in task performance was far more pronounced in tasks that were used to induce cognitive fatigue that were primarily visual (Ioannucci, 2022; Sun et al., 2022). Comparatively, induction tasks that were visually stimulating but heavily utilised other mechanisms to induce cognitive fatigue (such as the Stroop task or driving, which also requires a high level of decision-making and attention) did not always yield significant results from visual cognitive fatigue measures (Stroop, with Flanker task to measure; Habay et al., 2021; Van Cutsem et al., 2019).

As the Stroop has a visual aspect in order to trigger the inhibition and decision-making mechanisms, it seems that in order for it to be effective in measuring cognitive fatigue levels, a visual task is required as an inducing task. This was evidenced in Fortes and colleagues' (2021) study, whereby active videogame playing and smartphone social media interaction conditions for cognitive fatigue induction resulted in a slower response time in comparison to the control group (watching a sports coaching video). The primary difference between the videogame, smartphone and control groups were that the control group had a passive task that did not require any decision making. This indicates that the Stroop task was able to distinguish between cognitive fatigue levels, but only for the tasks that required similar resources (i.e., was visually stimulating and required decision making skills). Furthermore, as the Stroop can be used to induce cognitive fatigue (Ferreira et al., 2020; Habay et al., 2021; Van Cutsem et al., 2019), it is inappropriate to employ it in a real-world scenario.

Lastly, experiments two and three conducted by Ioannucci and colleagues (2022) used monocular visual stimuli saturation, which lasted for 41 minutes between the pre- and post-task measures. The measure utilised was the texture discrimination task, which is a visual task. The monocular visual stimuli saturation resulted in a significant decline for response accuracy for the texture discrimination task. However, post-hoc testing revealed that the decrease in performance accuracy was only revealed in the saturated eye condition. Comparatively, the non-saturated eye did not reveal a significant decrease in task performance accuracy when undergoing monocular viewing of the texture discrimination task. Therefore, it can be strongly suggested that the saturated eye was visually fatigued. These experiments consequently seem to evidence that the task performance of the texture discrimination is far more sensitive to visual fatigue change, and not sensitive to changes in overall objective cognitive fatigue.

The prominent indication from the studies included in this review suggests that modality specific tasks used for the measurement of objective cognitive fatigue can be impacted by the fatigue inducing task. Researchers should aim to consider how modality specific tasks could cause results that may be modality specific, and therefore cannot be attributed to a general measure of objective cognitive fatigue. For this reason, they should also not over-report significant results without discussing the limitations that modality specific tasks could incur.

Modality Specific Fatigue - Auditory

When an auditory task was used for induction of cognitive fatigue, participants performance within the induction task significantly decreased (as shown in the three auditory tasks used to induce cognitive fatigue in all the experiments by Ioannucci et al., 2022). This is to be expected, as the resource allocation for this modality was being specifically targeted. However, as mentioned, performance of the task used to induce cognitive fatigue is not a primary finding to be discussed, as we are exploring metrics that can be used to measure without the exacerbation of cognitive fatigue. When an auditory task was used as a measurement of cognitive fatigue for a separate induction task, the auditory tasks yielded mixed results that seemed to be dependent upon the induction task type. For example, when using real-world driving as a cognitive fatigue induction task, no significant differences were found in either Feyer and colleagues' (1994) or Riemersma and colleagues' (1978) studies. Similarly, no significant differences in task performance for reaction time, missed targets or false hits for the auditory oddball paradigm were found in Körber and colleagues' (2015) study, whereby simulated driving was used as an induction task. As driving, real-world or simulated, does not require a high level of auditory cognitive resource allocation, it is likely that, again, the use of auditory measures do not pick up on a decrease in general cognitive fatigue due to being modality specific. No articles included in this review used an auditory task for both the induction and measurement of cognitive fatigue within the same experimental setup. Therefore, further research into whether auditory measures perform better during modality specific investigations is required. However, researchers should again be aware of the limitations that modality specific measurement tasks can incur on results due to the interaction with the cognitively fatiguing tasks.

Therefore, the use of specific resources is an incredibly important factor to consider when addressing objective cognitive fatigue measurement. As discussed, modality specific induction tasks can cause modality specific results in terms of fatigue, and may not be appropriate for the overall assessment of an individuals' level of cognitive fatigue. Equally, attentional cognitive resources could become depleted and a measure that is specific to attentional skills may not be sensitive to these changes, causing important data to be accidentally omitted. A modality specific measure of objective cognitive fatigue may well be an appropriate step to take, if an individual is actually suffering from a modality specific type of fatigue due to their occupation, for example. However, this approach would require a large sample size of participants using individually appropriate objective measures of cognitive fatigue in order to assess the different categories of tasks utilised in a comparative way between individuals, which is less ideal in terms of robust research and widespread use of the measure for the broader population.

Task Performance - Reaction Time & Error Rate

A prominent use of reaction times and accuracy of task performance was utilised throughout many studies included in this review. A large part of the discussion surrounding reaction times and accuracy addresses vigilance and attentional levels, as well as the use of induction vs measurement tasks in terms of the specific use of attentional cognitive resources.

Reaction times were generally found to increase as time-on-task increased, however, task accuracy results were inconsistent, and often non-existent, following cognitive fatigue induction. Qin and colleagues (2021), Guillemin and colleagues (2022), Habay and colleagues (2021) and Van Cutsem and colleagues (2019) all found varying results. Following a simulated aviation task, Qin and colleagues (2021) recorded reaction times, missed trials and false starts (pre-emptive responses). The occurrence of false alarms and missed trials was so low that the authors did not statistically analyse the results. Reaction times did reveal significant increase, suggesting a decrease in vigilance over time. Similarly, in Van Cutsem and colleagues' (2019) study, accuracy was not affected by time-on-task, but participants reaction times became significantly increased. Comparatively, Guillemin and colleagues (2022) found both error rates and reaction times to increase as time-on-task increased in the high cognitive load condition of the TloadDback. Contrary to the other studies, the change in reaction times was not found to be significant. Habay and colleagues' (2021) study utilised a visuomotor task, Stroop task and Flanker task. Post-hoc testing revealed a significantly increased only for reaction times for the visuomotor task, however no significant changes in reaction times or accuracy were found for either the Stroop or Flanker tasks. Error rate of the advanced trail making task also did not show any significant difference between pre- and post- cognitive fatigue induction, however there was a numerical trend of increased error rate post- n-back task. This indicates that overall, error rate is not sensitive enough for the measurement of objective cognitive fatigue. Reaction times have a higher reliability, but still need to be approached with caution.

Where driving performance has been used separately as a measure of fatigue (where driving was not the fatigue inducing task), it is still questionable in its reliability due to the generically fatiguing aspect of the act of driving itself. Table 6 seems to display a relatively strong pattern of significant findings when using driving performance as a measure. When considered individually though, of the five studies that measured lane position deviation, only four reported whether their findings where significant, and only three of these studies found a significant increase in lane deviation. Speed maintenance was recorded in only two studies, with one study reporting a significant increase, and the other study showing an increase in speed in one of three conditions, however, significance was not reported. An increase in lane deviation and speed maintenance would be expected with an increase in cognitive fatigue due to decline in attentional resources, however, there is not enough evidence from the studies included in our systematic review alone to confidently suggest this is always the case. Furthermore, driving is also visually taxing, and therefore any non-physiological measures such as driving performance (for example, lane deviation or speed maintenance) could be correlated more significantly to changes in visual fatigue than cognitive fatigue. Regardless of this speculation, more literature needs to be reviewed before driving performance can be considered a reliable measure of objective cognitive fatigue. Furthermore, a task that can induce cognitive fatigue should not be considered as a useable measure of objective cognitive fatigue, and driving performance can ultimately only be measured due to the fatiguing act of driving itself.

Secondary tasks that are completely separate to driving performance may therefore be a more suitable alternative to indicate objective cognitive fatigue levels. Guo and colleagues' (2018) used a 90-minute simulated driving task, which resulted in the significant increase in reaction time and miss rate for Go/NoGo trials. However, no significant difference was found between the driving and control group (who watched a movie instead) in the percentage of false hits on the NoGo trials either pre- or post-fatigue manipulation. This seems to imply that decision making and inhibition for the Go/NoGo trials were not affected by either the driving or control group tasks. However, attention and vigilance for the Go/NoGo trials, determined by reaction time and miss rates, was significantly affected by the driving task and not the control task, indicating that cognitive resources were consumed more readily during the driving task in comparison to the control task. Similarly, the colour changing task, which was used as a secondary task to a real-world driving task in Riemersma and colleagues' (1978*a*) study, also showed a significant increase in missed and incorrect responses, and slower reaction times ('mental blocks'). Riese and colleagues (1999) also found that using a secondary dot counting task resulted in a decrease in the percentage of correct responses. This adds to the evidence base suggesting that driving results in decreased vigilance and attention, and therefore is a robust objective cognitive fatigue induction task.

It is worth noting however, that a limitation of using a visual based secondary task, such as dot counting or responding to a colour changing task, may be prone to reflecting modality specific visual fatigue as previously discussed rather than overall objective cognitive fatigue. Therefore, these tasks would need to be explored across other cognitive fatigue inducing tasks to help determine to what level the visual aspect of driving could be a confounding factor for objective cognitive fatigue measurement.

Several studies used work day tasks or similar tasks designed to simulate the same demand as a normal work day to explore this. Claros-Salinas and colleagues (2010) had participants complete tasks that vocationally reflected their individual work routines, and compared healthy control participants to stroke and multiple sclerosis patients. Both objective and subjective measures of cognitive fatigue suggested an increase, but task performance revealed no significant difference. However, these tasks only took place for one hour, at three time points throughout the day. In comparison, Völker and colleagues' (2016) study showed a decrease in reaction time following a work day that lasted between eight and nine hours. This discrepancy suggests that time-on-task has a pronounced effect on whether work day activities result in a realistic attentional decrease reflected by a decrease in reaction times.

Other studies that also had reaction time tasks to measure objective cognitive fatigue for work day related tasks (Jonasson et al., 2018; Zargari et al., 2018) had 40 and 30 minutes time-on-task respectively. Both found an unexpected improvement in task performance, which further indicates the importance of time-on-task being realistic of a eight hour work day. The performance improvement suggests that time-on-task used in both studies was not sufficient enough to overcome the initial effects of practice (Mangin et al., 2022). An alternative theory could be that a change in task could result in a refreshing redirection of attentional or modality resources, allowing for an increase in task performance (Behrens et al., 2023; Hopstaken et al., 2016). Despite an increase in task performance, subjective fatigue was still increased in Zargari and colleagues' (2018) study. Similarly, despite no significant differences in reaction time for Claros-Salinas and colleagues' (2010) study, subjective ratings of cognitive fatigue were also increased. This, firstly, reflects the disparity between subjective and objective cognitive fatigue and highlights the importance of implementing separate measures, and, secondly, could help to highlight the role of cognitive resource allocation with regards to subjective cognitive fatigue. It is possible that in some cases, feelings of cognitive fatigue could arise due to an overall decline in cognitive resources available, regardless of whether cognitive resources are more or less available for allocation depending on the task specifically at hand (for example, modality specific or goal-orientated). A lack of motivation for work day tasks, or having, for example, an occupation that requires significant lengths of screen time such as computer work, would readily impact the effortfulness required for allocation of resources in order to complete a given task. It is the effortfulness of the resource allocation that primarily gives rise to feelings of subjective cognitive fatigue, despite the potential of cognitive resources being available, which can lead to a 'normal' level of objective cognitive fatigue (i.e., no change in task performance) being simultaneous with an increase in subjective cognitive fatigue.

Fatigue Inducing

Combatting Fatigue Induction with Motivation

Younger individuals have more efficient use of cognitive resource consumption (Arnau et al., 2017; Zhang et al., 2023) and are less likely to have had previous or ongoing unknown conditions or illnesses that may have affected their individual cognitive fatigue resilience (Cohen et al., 2021). This would result in an expectation that younger individuals would be able to maintain a higher level of task performance, and therefore are less prone to objective cognitive fatigue. One study included in this review actively compared task performance between young and elderly participants (Zargari et al., 2018). This study found that overall task performance was significantly better in younger vs elderly participants, which fits this expectation. However, more recent research has indicated that, regardless of overall performance, task motivation within elderly individuals is managed more effectively than in younger individuals (Gilsoul et al., 2022; Wascher et al., 2016). This theory helps to bridge the gap as to why age does not always result in a significant decline in task performance, as motivation can help make up for increased consumption of available cognitive resources (Arnau et al., 2017; Terentjeviene et al., 2018). When considering Zargari and colleagues' (2018) results in more detail, this theory can be applied. Overall performance steadily increased with time-on-task in the elderly group, whereas the younger group had an initial improvement in task performance which then fluctuated as time-on-task increased.

Herlambang and colleagues (2019) also found time-on-task to be a prominent factor throughout their actively imposed high and low levels of intrinsic motivation conditions: a more pronounced decline in physiological results was revealed in the sudoku puzzles with low levels of intrinsic motivation. Similarly, performance measures declined over time, but more significantly in the low level of intrinsic motivation condition. These results are highly supportive of Figure 1, which was proposed in the introduction of this systematic review, whereby decreased availability of cognitive resources and decreased motivation results in objective cognitive fatigue, operationalised by a decrease in task performance.

Further evidence for the role of motivation is presented in Hopstaken and colleagues' (2016) study. Time-on-task was, again, the main factor for a decrease in task performance of the n-back task that was utilised in this study. However, in the final block of the n-back task, a reward manipulation was implemented. This significantly increased the accuracy of participants task performance. Reward manipulation results in an increase in motivation, which allows for individuals to overcome the effects of subjective feelings of cognitive fatigue (Kok, 2022) through the goal-directed allocation of resources. Both Herlambang and colleagues' (2019) and Hopstaken and colleagues' (2016) studies therefore strongly evidence the role of motivation in both increasing and decreasing objective cognitive fatigue levels.

Otto and colleagues (2018) found an interesting relationship between subjective and objective levels of cognitive fatigue in their study. Different n-back levels were used to manipulate cognitive workload, which resulted in significant differences in both error rates and the subjective rating scale of mental effort (RSME). However, reaction times did not show a significant effect with varying task load from the different n-back levels. An interesting result was also obtained from the subjective RSME results: only a slight increase in RSME score was found for the 3-back task condition, in comparison to a disproportionate increase in RSME score for the easier 1-back task condition. The authors concluded that both the RSME and reaction time results reflect disengagement due to boredom in the easier, 1-back task condition. As a result of this theory, the 3-back task was suggested to be more appealing for participants to actively engage in due to being more challenging. This deduction from the authors once again highlights the role of motivation as a linking factor throughout many of the studies included in this review. The authors are hypothesising that the lack of task engagement, or motivation to complete a task, affects the subjective levels of cognitive fatigue. This impacts the level of cognitive resource allocation that results in the decrease in task performance reflected in the increase of error rates. However, the contrasting result of the reaction times indicating that there was no effect of n-back condition on this aspect of task performance further indicates that a more fitting explanation for these results overall is a lack of motivation as the authors suggested, rather than a lack of available cognitive resources.

Furthermore, Magnuson and colleagues' (2021) study also used the n-back to cognitively fatigue participants. Participants' task performance did not show a significant change, although raw data did reveal an increase in reaction time and a decrease in response accuracy. EEG results did however, show a significant increase in alpha power, which indicates a decrease in alertness and attention levels (Gharagozlou et al., 2015). Despite the EEG results, the authors concluded that the 60 minutes spent completing the n-back task was not a long enough time-on-task period for cognitive fatigue induction. In terms of real-world use of the findings from this review, we can agree with the authors that, as EEG is not a field-usable measure, the task performance measures alone are not sensitive enough to register initial changes in objective cognitive fatigue levels for a 60-minute n-back task. However, a highly significant increase was found in the subjective ratings of cognitive fatigue and task aversion scores.

These studies also suggest that the n-back seems to be a difficult task to use for measuring cognitive fatigue objectively, as the higher load conditions further exacerbates objective cognitive fatigue in the process of the measurement, as shown by raw reaction time and accuracy data, and the lower condition n-back task induces boredom and lack of task engagement, which results in an increase in subjective cognitive fatigue. This is a problematic factor for the use of this task in real-world terms: people will not wish to engage in the task due to boredom, resulting in disengagement from daily usage of a measure. Consequently, the n-back can exacerbate objective cognitive fatigue both directly and indirectly through the impact of motivation and task engagement, which is not a desirable characteristic for the further development of objective cognitive fatigue measure.

Inhibition as Fatiguing

Literature indicates that cognitive fatigue is one of the most predominant factors in decreased response inhibition (Guo et al., 2018). The Go/NoGo tasks used in the studies included in this systematic review are the main task that requires a high level of response inhibition. Yu and colleagues (2022) used a Go/NoGo paradigm as a cognitive fatigue induction task, while Guo and colleagues (2018) and Claros-Salinas and colleagues (2010) used a Go/NoGo paradigm as a secondary task for measurement of cognitive fatigue. Yu and colleagues (2022) measured cognitive fatigue using a pupillometry. The authors subsequently found a significant increase in pupil diameter, however, this occurred only in Go trials, and not the NoGo trials which require the inhibitory decision-making response. Literature suggests that pupil diameter is correlated with cognitive load (Souchet et al., 2022), so it is possible that the fatiguing effects of inhibition are present only when a response is required.

Guo and colleagues' (2018) found no significant difference between the driving and control group in the percentage of false hits on the NoGo trials either pre- or post-fatigue manipulation. This is unexpected as the need for decision making resources are much higher in a driving scenario in comparison to the control condition. A possible explanation for this is that Guo and colleagues (2018) used a simulated driving task, rather than a real-world driving task. Therefore, a large component of normal decision making effort whilst driving is likely to have been removed to due to the safe nature of driving simulations. Using a Go/NoGo following a real-world driving condition may be more likely to yield significant results with regards to false hits as decision making resources may be more drastically consumed. A literature search for Go/NoGo paradigms on real-world driving, as opposed to simulated driving, revealed an obvious gap in research. This may be due to safety concerns.

Claros-Salinas and colleagues' (2010) study used a 3-minute Go/NoGo session for measurement of cognitive fatigue in stroke and multiple sclerosis (MS) patients, as well as a healthy control group, at three time-points throughout the day for two consecutive days. Reaction times of the Go/NoGo increased for both the patient groups over both days, and decreased over the two days for the control group. In terms of error rates, no significant effects were found for group main effect or for group x time-of-day interaction effects. The Go/NoGo task therefore seems to be heavily dependent on the fatiguing task employed, as two of the three studies included in this review had no significant changes in accuracy, which are indicative of inhibitory response control. Furthermore, it can be assumed that the decrease in cognitive resources from prolonged time-on-task can cause a consequent decrease in response inhibition, and should correlate with an increase in cognitive fatigue levels over time. Equally, response inhibition requires a high level of cognitive resources, meaning it can also cause cognitive fatigue. For this reason, Go/NoGo paradigms, or any other response inhibition task, are not suitable for the measurement of objective cognitive fatigue as they have the ability to simultaneously exacerbate it.

NON-INTRUSIVE

Most of the tasks have now already been discussed in line with the previous two criteria. However, if further or future research were to make arguments for any of the previously mentioned tasks, it is worth considering the fact that they are intrusive in the workplace. For example, many of the cognitive tasks require a large number of trials in order to get a sensitive, average measure, making them time-consuming (such as the Flanker task, reaction time tasks, Stroop task, n-back, etc.). It may therefore be inappropriate for an individual to carry out this task, particularly as it would likely need repeating several times a day. Even if work time adjustments were to be made, all of the above cognitive tasks, and in particular, auditory tasks, would require a quiet space to reduce confounding environmental factors, and a change in location or environmental noise levels may not always be a possibility.

This is a positive of physiological measures: many of them can be portable and left to passively collect data for long time periods, without the need for a location change, for example, ECG. However, this does not necessarily make the measure non-intrusive to the work day. As many factors can affect physiological measures, as previously discussed, differences in a typical work day between individuals (such as those in a busy healthcare setting vs desk occupations) will have hugely varying environmental and physiological factors, such as exercise and stress levels. Moreover, differences between an individual's day to day, such as working from home vs going into the office, or the varying emotional challenges that are encountered in a healthcare setting, will make comparisons challenging to verify on an individual level before even considering inter-individual comparisons.

Accessibility & Design Implementation

Currently, the only task that has met all the hierarchical criteria is the Critical Flicker Fusion task. Therefore, this task will be discussed in terms of all the criteria. The Critical Flicker Fusion task has been found to be relatively reliable in distinguishing between different levels of cognitive fatigue using the descending threshold. Additionally, it is a time efficient task that does not require a high level of cognitive resources, meaning that any potential fatiguing effects would be nominal (Mankowska et al., 2021). This is an ideal characteristic for measure development for a number of reasons; first, several trials could be conducted within a couple of minutes, providing an average, making the collected data more reliable. Secondly, the minimal amount of cognitive resources required allows for a significantly more robust comparison between individuals than an objective task performance measure, that would, to a greater extent, require cognitive resources. As a result of this, it is also unaffected by the practice effect, meaning accurate baselines can be obtained immediately. Thirdly, it is a relatively unobtrusive task because it is so time efficient, making it appropriate and easy for daily use in both personal and professional environments (Mankowska et al., 2021).

Furthermore, the Critical Flicker Fusion task is unlikely to exacerbate feelings of subjective cognitive fatigue in comparison to task performance measures, which would impact on the individuals' psyche and potentially exacerbate objective cognitive fatigue through the reallocation of cognitive resources. The setup of having a threshold also offers many advantages to this measure. Thresholds can, following future research with large sample sizes, be adjusted to differences between patient populations, allowing for comparison both within and between patient populations. This can be set against a general population threshold so that patients can understand and adjust to the impact that cognitive fatigue is having through a visual presentation of the threshold data, allowing them to make safe and informed decisions as to which daily or work day tasks they are capable of completing safely.

In terms of accessibility and design implementation, the Critical Flicker Fusion task has the potential to be widely available for a relatively small cost. It is also very easy to understand in terms of instruction, and should therefore be equally accessible to the majority of the general population. It is likely that, as many individuals carry on their person a mobile device which has a torch available, a mobile application can be made to simulate the Critical Flicker Fusion task utilising the torch function. This not only makes it accessible to anyone with a smartphone, but also allows it to be portable, adding to its ability to be used during a work day. Furthermore, data can be collected in a way that allows individuals to track their progress, and has the potential to be easily shared with healthcare professionals and colleagues remotely.

Despite all these advantages, the Critical Flicker Fusion task does have some minor limitations. For example, environmental lighting differences could potentially impact the threshold at which the participant can distinguish between flickers. Further research would need to be undertaken to determine the impact level this may have. However, the trade-off here is that the vast majority of other environmental factors are removed (i.e., noise levels and location). Moreover, where possible, measures can be taken by individuals to reduce the potential impact of lighting differences if they felt it was necessary.

A major limitation of the Critical Flicker Fusion task is that it is not suitable for those who have poor sight or are blind. This accounts for approximately 3% percent of the general adult population, and is expected to increase to approximately 5.4% percent due to demographic ageing (Pezzullo et al., 2018). There is also a high prevalence of new-onset visual loss or impairment following stroke (48 – 60%; Rowe et al., 2019). Recovery rates are varied but approximated at 30% for partial and 15% for full visual recovery, with high full recovery rates for specific types of visual impairments within the 15% statistic (Rowe et al., 2022). However, the prevalence of visual issues also increases with advancing age; the majority of stroke onsets were recorded in individuals aged 60 to 89 years, with a correspondingly high number of visual problems observed within this age bracket (Rowe et al., 2019). Therefore, the overall incidence of visual loss or impairment is not wholly applicable to the working age bracket of the stroke survivor population, meaning that the Critical Flicker Fusion task can still be considered a majority encompassing measure for return-to-work protocols.

CONCLUSION

As a result of in-depth interpretation of the results obtained from the studies included in this review, a prominent theme of motivation and resource-specific usage has come to light. Motivation plays a hugely important role in both the subjective and objective levels of cognitive fatigue – as feelings of subjective cognitive fatigue increase, the prioritisation of resource allocation shifts, resulting in a higher level of objective cognitive fatigue. Therefore, the use of reward and motivational incentives can help to overcome objective cognitive fatigue to a certain extent. It also provides a potential partial explanation towards the disparity that can be found between subjective and objective cognitive fatigue within an individual, as well as providing a potential explanation for individual differences in subjective and objective cognitive fatigue. This does however, highlight an important factor in terms of the tasks used throughout the studies: any task that measures objective performance has the ability to be impacted by motivation. If an individual is motivated by personal or social factors to return-to-work faster, and/or is determined to see a steady rate of improvement, motivation may skew the objective task performance to appear better than the individual's 'norm'.

Measuring the impact that an individual's work has on that level of cognitive fatigue is one of the root goals of conducting this systematic review. Because of the differing levels of cognitive fatigue that individuals will be subjected to during their work day, it is important to be able to determine cognitive fatigue levels that are personalised to the individual. This is because different processes and time on task will induce cognitive fatigue at different rates: it is important to not assume that two individuals carrying out the same tasks will have an equal level of cognitive fatigue by the end of the work day. Equally it is important not to assume that an individual has a higher level of cognitive fatigue because they have a seemingly more taxing work day in comparison to another who may be affected more by a seemingly less taxing work day. Being able to measure cognitive fatigue levels objectively will require the removal of bias within social constructs regarding what is believed to be a taxing work day and base it solely on an individual's needs. This bias needs to be removed from both the individual and what they think should fatigue them - which may affect subjective levels of cognitive fatigue - as well as from encroaching expectations from colleagues surrounding them.

This highlights another significant advantage of the Critical Flicker Fusion task: as long as the individual gives an honest answer, it cannot be skewed by motivation or subjective feelings of cognitive fatigue. As such, future research should consider the application of this task as a widespread measure of objective cognitive fatigue based on its reliability, accessibility, and straight-forward, non-fatiguing advantages that allow for collaboration, individualised thresholds, and wider threshold comparisons. This could be achieved by focusing on the improvement and implementation of smartphone applications that can reliably perform the Critical Flicker Fusion task. A smartphone application, "Critical Flicker Fusion" by MukhinApps, is already freely available for Android users which utilises the screen display. Furthermore, Zarei and colleagues (2017) presented a novel smartphone application (eyeFusion) which can objectively measure the critical flicker fusion threshold of an individual using flickering stimuli on the screen display. Applications such as these should be tested on a large sample size to determine their reliability and validity. Any improvements to the application can then be completed, if required, before gathering larger data samples for analysis. The analysis should focus on finding normalised thresholds which can be used as a reference point and subsequently allow for comparative analysis for those who may be prone to higher levels of objective cognitive fatigue. The use of a free smartphone app is the epitome of accessibility for large populations on a daily basis for the measurement and long-term tracking of objective cognitive fatigue.

Another way the critical flicker fusion task could be implemented is through clinicians. A specific, portable flicker apparatus (akin to a blood pressure machine, for example) could be distributed for real-time observation of objective cognitive fatigue during appointments. This would be a fast and easy implementation for clinicians as the device would require very little training and time to complete. Moreover, data could be recorded as a threshold number in an individuals' medical file, making it easily trackable over time. This could lead to potential diagnoses of cognitive fatigue conditions, such as myalgic encephalomyelitis, or tracking improvement or decline of objective cognitive fatigue following a brain injury or stroke. This would be ideal in helping to inform clinicians of an individuals' ability to return to work, by repeatedly reaching an acceptable threshold following normal daily activity. This would still need to be coupled with a questionnaire regarding the individuals' subjective level of cognitive fatigue and a general overview of the days' activities. However, in combination with a smartphone app that would allow for more data points over a longer time period, a more conclusive overall picture of an individuals' objective cognitive fatigue levels could be obtained.

Limitations of the Review

It is notable that, despite strict inclusion criteria, there were an incredible number of differences between all the studies which made comparisons challenging. This was made additionally complex due to the variation in referring to and defining cognitive fatigue. Whilst not an initial objective of the research, upon reading the literature there were a number of terms that were used interchangeably to imply cognitive fatigue and a number of different ways of defining cognitive fatigue. As such, we opted to synthesise the definitions of cognitive fatigue across the included studies. The most common definition of objective cognitive fatigue when it was provided throughout the articles included in this systematic review comprise of mentioning, as a minimum, a decrease in task performance due to a decline in available cognitive resources. General definitions of cognitive fatigue additionally included the subjective constituents of feelings of tiredness, discomfort and lethargy, and an unwillingness to complete a task (i.e., low motivation) due to effortfulness.

Furthermore, some studies were excluded on a reliability basis due to the exclusion criteria requiring more than one objective measure of cognitive fatigue in order to cross-validate the data. Further research could be undertaken to determine if studies with only one objective measure could be re-analysed alongside a previously verified objective measure to assess reliability.

Lastly, although a meta-analysis was going to be undertaken, a large proportion of studies did not provide effect sizes for their results. A total of fifteen studies reported effect sizes, with three studies reporting effect sizes, but not for all measures. A huge 38 studies out of the 56 included in this review did not report any effect sizes. This highlights the lack of open science practices in reporting effect sizes alongside inferential statistics which should be correctly implemented going forwards.

References

Abdelhamid, G. S. M., Bassiouni, M. G. A., & Gómez-Benito, J. (2021). Assessing Cognitive Abilities Using the WAIS-IV: An Item Response Theory Approach. *International Journal of Environmental Research and Public Health*, *18*(13), 6835. https://doi.org/10.3390/ijerph18136835

Abdullah, N., & Mustapar, N. A. (2019). *Heart rate reading patterns at moment of information overload during online information searching*. *3*.

Adalberto Ferreira Junior, Chierotti, P., Gabardo, J. M., Giovanini, B., Okano, A. H., Buzzachera, C. F., Okazaki, V. H. A., Okuno, N. M., & Altimari, L. R. (2020). Efectos residuales de la fatiga mental en la fatiga subjetiva, tiempo de reacción y respuestas cardíacas / Residual effects of mental fatigue on subjective fatigue, reaction time and cardiac responses. *Revista de Psicología Del Deporte*, *29*(2), 27–34. Publicly Available Content Database.

Ahmed, S. K., Mohammed, M. G., Abdulqadir, S. O., El-Kader, R. G. A., El-Shall, N. A., Chandran, D., Rehman, M. E. U., & Dhama, K. (2023). Road traffic accidental injuries and deaths: A neglected global health issue. *Health Science Reports*, *6*(5), e1240. https://doi.org/10.1002/hsr2.1240

Álvaro, J. L., Garrido, A., Pereira, C. R., Torres, A. R., & Barros, S. C. (2019). Unemployment, Self-esteem, and Depression: Differences between Men and Women. *The Spanish Journal of Psychology*, *22*, E1. https://doi.org/10.1017/sjp.2018.68

Arnau, S., Möckel, T., Rinkenauer, G., & Wascher, E. (2017). The interconnection of mental fatigue and aging: An EEG study. *International Journal of Psychophysiology*, *117*, 17–25. https://doi.org/10.1016/j.ijpsycho.2017.04.003

Aromataris, E., Fernandez, R., Godfrey, C., Holly, C., Khalil, H., & Tungpunkom, P. (2020). Chapter 10: Umbrella reviews. In *JBI Manual for Evidence Synthesis*. JBI. https://doi.org/10.46658/JBIMES-20-11

Ashman, T. A., Cantor, J. B., Gordon, W. A., Spielman, L., Egan, M., Ginsberg, A., Engmann, C., Dijkers, M., & Flanagan, S. (2008). Objective Measurement of Fatigue Following Traumatic Brain Injury. *Journal of Head Trauma Rehabilitation*, *23*(1), 33–40. https://doi.org/10.1097/01.HTR.0000308719.70288.22

Assecondi, S., Villa-Sánchez, B., & Shapiro, K. (2022). Event-Related Potentials as Markers of Efficacy for Combined Working Memory Training and Transcranial Direct Current Stimulation Regimens: A Proof-of-Concept Study. *Frontiers in Systems Neuroscience*, *16*. https://www.frontiersin.org/articles/10.3389/fnsys.2022.837979

Bafna, T., & Hansen, J. P. (2021). Mental fatigue measurement using eye metrics: A systematic literature review. *Psychophysiology*, *58*(6), e13828. https://doi.org/10.1111/psyp.13828

Balasubramanian, V., & Bhardwaj, R. (2018). Can cECG be an unobtrusive surrogate to determine cognitive state of driver? *Transportation Research Part F: Traffic Psychology and Behaviour*, *58*, 797–806. https://doi.org/10.1016/j.trf.2018.07.011

Baldassarre, G. (2011). What are intrinsic motivations? A biological perspective. *Proc. Int. Conf. Develop. Learn. Epigenetic Robot. (ICDL-EpiRob-2011)*, E1–E8.

Barboza, M. A., Becerra, L. C., Serrano, F. E., & Arauz, A. (2017). Abstract WP162: Work Performance Questionnaire After Returning to Work in Ischemic Stroke Survivors: The Work Performance Survey. *Stroke*, *48*(suppl_1), AWP162–AWP162. https://doi.org/10.1161/str.48.suppl_1.wp162

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276–292. https://doi.org/10.1037/0033-2909.91.2.276

Behrens, M., Gube, M., Chaabene, H., Prieske, O., Zenon, A., Broscheid, K.-C., Schega, L., Husmann, F., & Weippert, M. (2023). Fatigue and Human Performance: An Updated Framework. *Sports Medicine*, *53*(1), 7–31. https://doi.org/10.1007/s40279-022-01748-2

Benoit, C.-E., Solopchuk, O., Borragán, G., Carbonnelle, A., Durme, S. V., & Zénon, A. (2017). Objective but not subjective fatigue increases cognitive task avoidance. *bioRxiv*, 208322. https://doi.org/10.1101/208322

Billones, R. R., Kumar, S., & Saligan, L. N. (2020). Disentangling fatigue from anhedonia: A scoping review. *Translational Psychiatry*, *10*(1), 273. https://doi.org/10.1038/s41398-020-00960-w

Borragán, G., Slama, H., Bartolomei, M., & Peigneux, P. (2017). Cognitive fatigue: A Time-based Resource-sharing account. *Cortex*, *89*, 71–84. https://doi.org/10.1016/j.cortex.2017.01.023

Chaudhuri, A., & Behan, P. O. (2004). Fatigue in neurological disorders. *The Lancet*, *363*(9413), 978–988. https://doi.org/10.1016/S0140-6736(04)15794-2

Chen, Y., Fang, W., Guo, B., & Bao, H. (2021). Fatigue-Related Effects in the Process of Task Interruption on Working Memory. *Frontiers in Human Neuroscience*. Publicly Available Content Database. https://doi.org/10.3389/fnhum.2021.703422

Christodoulou, C., Schneider, S., Junghaenel, D. U., Broderick, J. E., & Stone, A. A. (2014). Measuring daily fatigue using a brief scale adapted from the Patient-Reported Outcomes Measurement Information System (PROMIS®). *Quality of Life Research*, *23*(4), 1245–1253. https://doi.org/10.1007/s11136-013-0553-z

Claros-Salinas, D., Bratzke, D., Greitemann, G., Nickisch, N., Ochs, L., & Schröter, H. (2010). Fatigue-related diurnal variations of cognitive performance in multiple sclerosis and stroke patients. *Journal of the Neurological Sciences*, *295*(1–2), 75–81. https://doi.org/10.1016/j.jns.2010.04.018

Cohen, R. W., Meinhardt, A. J., Gmelin, T., Qiao, Y. (Susanna), Moored, K. D., Katz, R. D., Renner, S. W., & Glynn, N. W. (2021). Prevalence and severity of perceived mental fatigability in older adults: The Long Life Family Study. *Journal of the American Geriatrics Society*, *69*(5), 1401–1403. https://doi.org/10.1111/jgs.17075

Cook, D. B., O'Connor, P. J., Lange, G., & Steffener, J. (2007). Functional neuroimaging correlates of mental fatigue induced by cognition among chronic fatigue syndrome patients and controls. *NeuroImage*, *36*(1), 108–122. https://doi.org/10.1016/j.neuroimage.2007.02.033

Coole, C., Radford, K., Grant, M., & Terry, J. (2013). Returning to Work After Stroke: Perspectives of Employer Stakeholders, a Qualitative Study. *Journal of Occupational Rehabilitation*, *23*(3), 406–418. https://doi.org/10.1007/s10926-012-9401-1

Dantzer, R., Heijnen, C. J., Kavelaars, A., Laye, S., & Capuron, L. (2014). The neuroimmune basis of fatigue. *Trends in Neurosciences*, *37*(1), 39–46. https://doi.org/10.1016/j.tins.2013.10.003

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. Academic Press.

Dawson, D., & Fletcher, A. (2001). A quantitative model of work-related fatigue: Background and definition. *Ergonomics*, *44*(2), 144–163. https://doi.org/10.1080/00140130119399

Dawson, D., Ian Noy, Y., Härmä, M., Åkerstedt, T., & Belenky, G. (2011). Modelling fatigue and the use of fatigue models in work settings. *Accident Analysis & Prevention*, *43*(2), 549–564. https://doi.org/10.1016/j.aap.2009.12.030

de Waard, D., & Brookhuis, K. A. (1991). Assessing driver status: A demonstration experiment on the road. *Accident Analysis & Prevention*, *23*(4), 297–307. https://doi.org/10.1016/0001-4575(91)90007-R

DeLuca, J., Genova, H. M., Capili, E. J., & Wylie, G. R. (2009). Functional neuroimaging of fatigue. *Physical Medicine and Rehabilitation Clinics of North America*, *20*(2), 325–337. https://doi.org/10.1016/j.pmr.2008.12.007

Dickens, A., Schenke, K., Champion, A., & Edgar, G. (2022). *To what extent can current tools objective measure cognitive fatigue?* PROSPERO. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=323292

Ferreira, A., Chierotti, P., Gabardo, J. M., Giovanini, B., Okano, A. H., Buzzachera, C. F., Okazaki, V. H. A., Okuno, N. M., & Altimari, L. R. (2020). *Residual Effects of Mental Fatigue on Subjective Fatigue, Reaction Time and Cardiac Responses*.

Feyer, A.-M., Williamson, A. M., & Friswell, R. (1994). *Strategies to combat fatigue in the long distance road transport industry: Stage 2, Evaluation of alternative work practices.* https://trid.trb.org/View/550470

Gartenberg, D., Gunzelmann, G., Hassanzadeh-Behbaha, S., & Trafton, J. G. (2018). Examining the Role of Task Requirements in the Magnitude of the Vigilance Decrement. *Frontiers in Psychology*, *9*, 1504. https://doi.org/10.3389/fpsyg.2018.01504

Gatti, E., Calzolari, E., Maggioni, E., & Obrist, M. (2018). Emotional ratings and skin conductance response to visual, auditory and haptic stimuli. *Scientific Data*, *5*(1), Article 1. https://doi.org/10.1038/sdata.2018.120

Genova, H. M., Rajagopalan, V., DeLuca, J., Das, A., Binder, A., Arjunan, A., Chiaravalloti, N., & Wylie, G. (2013). Examination of Cognitive Fatigue in Multiple Sclerosis using

Functional Magnetic Resonance Imaging and Diffusion Tensor Imaging. *PLoS ONE*, *8*(11), e78811. https://doi.org/10.1371/journal.pone.0078811

Gergelyfi, M., Jacob, B., Olivier, E., & Zénon, A. (2015). Dissociation between mental fatigue and motivational state during prolonged mental activity. *Frontiers in Behavioral Neuroscience*, *9*. https://doi.org/10.3389/fnbeh.2015.00176

Gharagozlou, F., Saraji, G. N., Mazloumi, A., Nahvi, A., Nasrabadi, A. M., Foroushani, A. R., Kheradmand, A. A., Ashouri, M., & Samavati, M. (2015). Detecting Driver Mental Fatigue Based on EEG Alpha Power Changes during Simulated Driving. *Iranian Journal of Public Health*, *44*(12), 1693–1700.

Gilsoul, J., Libertiaux, V., & Collette, F. (2022). Cognitive fatigue in young, middle-aged, and older: Breaks as a way to recover. *Applied Psychology*, *71*(4), 1565–1597. https://doi.org/10.1111/apps.12358

Giuliano, R. J., Karns, C. M., Bell, T. A., Petersen, S., Skowron, E. A., Neville, H. J., & Pakulak, E. (2018). Parasympathetic and sympathetic activity are associated with individual differences in neural indices of selective attention in adults. *Psychophysiology*, *55*(8), e13079. https://doi.org/10.1111/psyp.13079

Green, T., McGovern, H. T., & Hinkle, J. (2021). Understanding Return to Work After Stroke Internationally: A Scoping Review. *Journal of Neuroscience Nursing*, *Publish Ahead of Print*. https://doi.org/10.1097/JNN.0000000000000603

Greenberg, D. B. (2002). Clinical Dimensions of Fatigue. *Primary Care Companion to The Journal of Clinical Psychiatry*, *4*(3), 90–93.

Guillemin, C., Hammad, G., Read, J., Requier, F., Charonitis, M., Delrue, G., Vandeleene, N., Lommers, E., Maquet, P., & Collette, F. (2022). Pupil response speed as a marker of cognitive fatigue in early Multiple Sclerosis☆. *Multiple Sclerosis and Related Disorders*, *65*, 104001. https://doi.org/10.1016/j.msard.2022.104001

Guo, Z., Chen, R., Liu, X., Zhao, G., Zheng, Y., Gong, M., & Zhang, J. (2018). The impairing effects of mental fatigue on response inhibition: An ERP study. *PLOS ONE*, *13*(6), e0198206. https://doi.org/10.1371/journal.pone.0198206

Guo, Z., Chen, R., Zhang, K., Pan, Y., & Wu, J. (2016). The Impairing Effect of Mental Fatigue on Visual Sustained Attention under Monotonous Multi-Object Visual Attention Task in Long Durations: An Event-Related Potential Based Study. *PloS One*, *11*(9), e0163360. https://doi.org/10.1371/journal.pone.0163360

Habay, J., Proost, M., De Wachter, J., Díaz-García, J., De Pauw, K., Meeusen, R., Jeroen Van Cutsem, & Roelands, B. (2021). Mental Fatigue-Associated Decrease in Table Tennis Performance: Is There an Electrophysiological Signature? *International Journal of Environmental Research and Public Health*, *18*(24), 12906. Publicly Available Content Database. https://doi.org/10.3390/ijerph182412906

Harrington, M. E. (2012). Neurobiological studies of fatigue. *Progress in Neurobiology*, *99*(2), 93–105. https://doi.org/10.1016/j.pneurobio.2012.07.004

Hartmann, K. E., Matchar, D. B., & Chang, S. M. (2012, June). *Using the PICOTS framework to assess and describe applicability of medical tests* [Text]. Agency for Healthcare Research and Quality (US). https://www.ncbi.nlm.nih.gov/books/NBK98234/table/ch6.t1/

Herlambang, M. B., Cnossen, F., & Taatgen, N. A. (2021). The effects of intrinsic motivation on mental fatigue. *PLoS One*, *16*(1). Publicly Available Content Database. https://doi.org/10.1371/journal.pone.0243754

Herlambang, M. B., Taatgen, N. A., & Cnossen, F. (2019). The Role of Motivation as a Factor in Mental Fatigue. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *61*(7), 1171–1185. https://doi.org/10.1177/0018720819828569

Hiploylee, C., Dufort, P. A., Davis, H. S., Wennberg, R. A., Tartaglia, M. C., Mikulis, D., Hazrati, L.-N., & Tator, C. H. (2017). Longitudinal Study of Postconcussion Syndrome: Not Everyone Recovers. *Journal of Neurotrauma*, *34*(8), 1511–1523. https://doi.org/10.1089/neu.2016.4677

Holgado, D., Sanabria, D., Perales, J. C., & Vadillo, M. A. (2020). Mental Fatigue Might Be Not So Bad for Exercise Performance After All: A Systematic Review and Bias-Sensitive Meta-Analysis. *Journal of Cognition*, *3*(1), 38. https://doi.org/10.5334/joc.126

Holtzer, R., Shuman, M., Mahoney, J. R., Lipton, R., & Verghese, J. (2011). Cognitive Fatigue Defined in the Context of Attention Networks. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *18*(1), 108–128. https://doi.org/10.1080/13825585.2010.517826

Hopstaken, J. (2016). *Conquering fatigue: The battle for engagement*. Erasmus University Rotterdam.

Hopstaken, J. F., van der Linden, D., Bakker, A. B., & Kompier, M. A. J. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement: Mental fatigue and task disengagement. *Psychophysiology*, *52*(3), 305–315. https://doi.org/10.1111/psyp.12339

Hopstaken, J. F., van der Linden, D., Bakker, A. B., Kompier, M. A. J., & Leung, Y. K. (2016). Shifts in attention during mental fatigue: Evidence from subjective, behavioral, physiological, and eye-tracking data. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 878–889. https://doi.org/10.1037/xhp0000189

Hopstaken, J. F., Wanmaker, S., van der Linden, D., & Bakker, A. B. (2015). Does Dysphoria Lead to Divergent Mental Fatigue Effects on a Cognitive Task? *PLoS One*, *10*(6). Publicly Available Content Database. https://doi.org/10.1371/journal.pone.0130304

Hu, X., & Lodewijks, G. (2020). Detecting fatigue in car drivers and aircraft pilots by using non-invasive measures: The value of differentiation of sleepiness and mental fatigue. *Journal of Safety Research*, *72*, 173–187. https://doi.org/10.1016/j.jsr.2019.12.015

Iancheva, D., Trenova, A., Mantarova, S., & Terziyski, K. (2019). Functional Magnetic Resonance Imaging Correlations Between Fatigue and Cognitive Performance in Patients With Relapsing Remitting Multiple Sclerosis. *Frontiers in Psychiatry*, *10*, 754. https://doi.org/10.3389/fpsyt.2019.00754

Investigations | Diagnosis | Tiredness/fatigue in adults | CKS | NICE. (n.d.). Retrieved 14 December 2021, from https://cks.nice.org.uk/topics/tiredness-fatigue-in-adults/diagnosis/investigations/

Ioannucci, S. (2022). Exploring of the the neural mechanisms of fatigue by passive induction. *Thesis*.

Ioannucci, S., Borragán, G., & Zénon, A. (2022). Passive visual stimulation induces fatigue under conditions of high arousal elicited by auditory tasks. *Journal of Experimental Psychology: General*, *151*(12), 3097–3113. https://doi.org/10.1037/xge0001224

Jagannath, M., & Balasubramanian, V. (2014). Assessment of early onset of driver fatigue using multimodal fatigue measures in a static simulator. *Applied Ergonomics*, *45*(4), 1140–1147. https://doi.org/10.1016/j.apergo.2014.02.001

Jimenez-Pinto, J., & Torres-Torriti, M. (2013). Optical Flow and Driver's Kinematics Analysis for State of Alert Sensing. *Sensors*, *13*(4), 4225–4257. Publicly Available Content Database. https://doi.org/10.3390/s130404225

Joanna Briggs Institute. (2014). *JBI Critical Appraisal Tools | JBI*. https://jbi.global/critical-appraisal-tools

Johansson, B., & Rönnbäck, L. (2012). Mental fatigue and cognitive impairment after an almost neurological recovered stroke. *ISRN Psychiatry*, *2012*, 686425. https://doi.org/10.5402/2012/686425

Jonasson, A., Levin, C., Renfors, M., Strandberg, S., & Johansson, B. (2018). Mental fatigue and impaired cognitive function after an acquired brain injury. *Brain and Behavior*, *8*(8). Publicly Available Content Database. https://doi.org/10.1002/brb3.1056

Kim, H.-J., Park, H.-K., Lim, D.-W., Choi, M.-H., Kim, H.-J., Lee, I.-H., Kim, H.-S., Choi, J.-S., Tack, G.-R., & Chung, S.-C. (2013). Effects of oxygen concentration and flow rate on cognitive ability and physiological responses in the elderly. *Neural Regeneration Research*, *8*(3), 264–269. https://doi.org/10.3969/j.issn.1673-5374.2013.03.009

Kisala, P. A., Bushnik, T., Boulton, A. J., Hanks, R. A., Kolakowsky-Hayner, S. A., & Tulsky, D. S. (2019). Measuring Fatigue in TBI: Development of the TBI-QOL Fatigue Item Bank and Short Form. *Journal of Head Trauma Rehabilitation*, *34*(5), 289–297. https://doi.org/10.1097/HTR.0000000000000530

Ko, L.-W., Lai, W.-K., Liang, W.-G., Chuang, C.-H., Lu, S.-W., Lu, Y.-C., Hsiung, T.-Y., Wu, H.-H., & Lin, C.-T. (2015). Single channel wireless EEG device for real-time fatigue level detection. *2015 International Joint Conference on Neural Networks (IJCNN)*, 1–5. https://doi.org/10.1109/IJCNN.2015.7280817

Kobayashi, F., Watanabe, T., Watanabe, M., Akamatsu, Y., Tomita, T., Nakane, T., Furui, H., Takeuchi, K., Okada, A., Ohashi, R., & Hayano, J. (2002). Blood Pressure and Heart Rate Variability in Taxi Drivers on Long Duty Schedules. *Journal of Occupational Health*, *44*(4), 214–220. https://doi.org/10.1539/joh.44.214

Kok, A. (2022). Cognitive control, motivation and fatigue: A cognitive neuroscience perspective. *Brain and Cognition*, *160*, 105880. https://doi.org/10.1016/j.bandc.2022.105880

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

Laine, C. M., Spitler, K. M., Mosher, C. P., & Gothard, K. M. (2009). Behavioral Triggers of Skin Conductance Responses and Their Neural Correlates in the Primate Amygdala. *Journal of Neurophysiology*, *101*(4), 1749–1754. https://doi.org/10.1152/jn.91110.2008

Lange, G., Steffener, J., Cook, D., Bly, B., Christodoulou, C., Liu, W., Deluca, J., & Natelson, B. (2005). Objective evidence of cognitive complaints in Chronic Fatigue Syndrome: A BOLD fMRI study of verbal working memory. *NeuroImage*, *26*(2), 513–524. https://doi.org/10.1016/j.neuroimage.2005.02.011

LeDuc, P., Greig, J., & Dumond, S. (2005). *Self-Report and Ocular Measures of Fatigue in U.S. Army Apache Aviators Following Flight*. https://www.semanticscholar.org/paper/Self-Report-and-Ocular-Measures-of-Fatigue-in-U.S.-LeDuc-Greig/8a8529433d26741f8ef333d1a205619a75a359a6

Li, Z., Jiao, K., Chen, M., & Wang, C. (2004). Reducing the effects of driving fatigue with magnitopuncture stimulation. *Accident; Analysis and Prevention*, *36*(4), 501-505. https://doi.org/10.1016/S0001-4575(03)00044-7

Lin, F. V., Tao, Y., Chen, Q., Anthony, M., Zhang, Z., Tadin, D., & Heffner, K. L. (2020). Processing speed and attention training modifies autonomic flexibility: A mechanistic intervention study. *NeuroImage*, *213*, 116730. https://doi.org/10.1016/j.neuroimage.2020.116730

Liu, Y.-C., & Wu, T.-J. (2008). Fatigued driver's driving behavior and cognitive task performance: Effects of road environments and road environment changes. *Safety Science*, *47*(8), 1083–1089. https://doi.org/10.1016/j.ssci.2008.11.009

Losoi, H., Silverberg, N. D., Wäljas, M., Turunen, S., Rosti-Otajärvi, E., Helminen, M., Luoto, T. M., Julkunen, J., Öhman, J., & Iverson, G. L. (2016). Recovery from Mild Traumatic Brain Injury in Previously Healthy Adults. *Journal of Neurotrauma*, *33*(8), 766–776. https://doi.org/10.1089/neu.2015.4070

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

Magnuson, J. R., Doesburg, S. M., & McNeil, C. J. (2021). Development and recovery time of mental fatigue and its impact on motor function. *Biological Psychology*, *161*, 108076. https://doi.org/10.1016/j.biopsycho.2021.108076

Mangin, T., Audiffren, M., Lorcery, A., Mirabelli, F., Benraiss, A., & André, N. (2022). A plausible link between the time-on-task effect and the sequential task effect. *Frontiers in Psychology*, *13*, 998393. https://doi.org/10.3389/fpsyg.2022.998393

Manjaly, Z.-M., Harrison, N. A., Critchley, H. D., Do, C. T., Stefanics, G., Wenderoth, N., Lutterotti, A., Müller, A., & Stephan, K. E. (2019). Pathophysiological and cognitive mechanisms of fatigue in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*, *90*(6), 642–651. https://doi.org/10.1136/jnnp-2018-320050

Mankowska, N. D., Marcinkowska, A. B., Waskow, M., Sharma, R. I., Kot, J., & Winklewski, P. J. (2021). Critical Flicker Fusion Frequency: A Narrative Review. *Medicina*, *57*(10), Article 10. https://doi.org/10.3390/medicina57101096

Manly, T. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia*, *37*(6), 661–670. https://doi.org/10.1016/S0028-3932(98)00127-4

Massar, S. A. A., Csathó, Á., & Van der Linden, D. (2018). Quantifying the Motivational Effects of Cognitive Fatigue Through Effort-Based Decision Making. *Frontiers in Psychology*, *9*, 843. https://doi.org/10.3389/fpsyg.2018.00843

Massar, S., Wester, A., Volkerts, E., & Kenemans, J. (2010). Manipulation specific effects of mental fatigue: Evidence from novelty processing and simulated driving. *Psychophysiology*, *47*, 1119–1126. https://doi.org/10.1111/j.1469-8986.2010.01028.x

Matsuo, T., Ishii, A., & Yoshikawa, T. (2022). Neural effects of viewing children's faces on mental fatigue: A magnetoencephalography study. *Experimental Brain Research*, *240*(11), 2885–2896. https://doi.org/10.1007/s00221-022-06466-9

Matuz, A., van der Linden, D., Kisander, Z., Hernádi, I., Kázmér, K., & Csathó, Á. (2021). Enhanced cardiac vagal tone in mental fatigue: Analysis of heart rate variability in Time-on-Task, recovery, and reactivity. *PLoS ONE*, *16*(3), e0238670. https://doi.org/10.1371/journal.pone.0238670

McKinley, R. A., McIntire, L. K., Schmidt, R., Repperger, D. W., & Caldwell, J. A. (2011). Evaluation of eye metrics as a detector of fatigue. *Human Factors*, *53*(4), 403–414. https://doi.org/10.1177/0018720811411297

Mehmood, I., Li, H., Umer, W., Arsalan, A., Saad Shakeel, M., & Anwer, S. (2022). Validity of facial features' geometric measurements for real-time assessment of mental fatigue in construction equipment operators. *Advanced Engineering Informatics*, *54*, 101777. https://doi.org/10.1016/j.aei.2022.101777

Mendoza-Ruvalcaba, N. M., von Humboldt, S., Flores-Villavicencio, M. E., & Díaz-García, I. F. (2019). Psychological Fatigue. In D. Gu & M. E. Dupre (Eds.), *Encyclopedia of Gerontology and Population Aging* (pp. 1–5). Springer International Publishing. https://doi.org/10.1007/978-3-319-69892-2_85-1

Mergenthaler, P., Lindauer, U., Dienel, G. A., & Meisel, A. (2013). Sugar for the brain: The role of glucose in physiological and pathological brain function. *Trends in Neurosciences*, *36*(10), 587–597. https://doi.org/10.1016/j.tins.2013.07.001

Mink, J. W., Blumenschine, R. J., & Adams, D. B. (1981). Ratio of central nervous system to body metabolism in vertebrates: Its constancy and functional basis. *The American Journal of Physiology*, *241*(3), R203-212. https://doi.org/10.1152/ajpregu.1981.241.3.R203

Mizuno, K., Tajima, K., Watanabe, Y., & Kuratsune, H. (2014). Fatigue correlates with the decrease in parasympathetic sinus modulation induced by a cognitive challenge. *Behavioral and Brain Functions*, *10*, 25. Publicly Available Content Database. https://doi.org/10.1186/1744-9081-10-25

Mizuno, K., Tanaka, M., Tajima, K., Okada, N., Rokushima, K., & Watanabe, Y. (2010). Effects of mild-stream bathing on recovery from mental fatigue. *Medical Science Monitor*, *16*(1), CR8-14.

Mizuno, K., Tanaka, M., Tanabe, H. C., Joudoi, T., Kawatani, J., Shigihara, Y., Tomoda, A., Miike, T., Imai-Matsumura, K., Sadato, N., & Watanabe, Y. (2015). Less efficient and costly processes of frontal cortex in childhood chronic fatigue syndrome. *NeuroImage: Clinical*, *9*, 355–368. https://doi.org/10.1016/j.nicl.2015.09.001

Mizuno, K., Tanaka, M., Yamaguti, K., Kajimoto, O., Kuratsune, H., & Watanabe, Y. (2011). Mental fatigue caused by prolonged cognitive load associated with sympathetic hyperactivity. *Behavioral and Brain Functions*, *7*(1), 17. https://doi.org/10.1186/1744-9081-7-17

Möckel, T., Beste, C., & Wascher, E. (2015). The Effects of Time on Task in Response Selection—An ERP Study of Mental Fatigue. *Scientific Reports*, *5*(1), 10113. https://doi.org/10.1038/srep10113

Moradi, A., Nazari, S. S. H., & Rahmani, K. (2019). Sleepiness and the risk of road traffic accidents: A systematic review and meta-analysis of previous studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *65*, 620–629. https://doi.org/10.1016/j.trf.2018.09.013

MukhinApps. (2024). *Critical Flicker Fusion* (Version 1.2) [Mobile app]. Google Play Store. https://apkpure.com/critical-flicker-fusion/com.vm.CFF

Mulder, G., & Mulder, L. J. M. (1981). Information Processing and Cardiovascular Control. *Psychophysiology*, *18*(4), 392–402. https://doi.org/10.1111/j.1469-8986.1981.tb02470.x

Muñoz-de-Escalona, E., Cañas, J. J., & Noriega, P. (2020). Inconsistencies between mental fatigue measures under compensatory control theories. *Psicológica Journal*, *41*(2), 103–126. https://doi.org/10.2478/psicolj-2020-0006

Naeeri, S., Kang, Z., Mandal, S., & Kim, K. (2021). Multimodal Analysis of Eye Movements and Fatigue in a Simulated Glass Cockpit Environment. *Aerospace*, *8*(10), Article 10. https://doi.org/10.3390/aerospace8100283

Norder, G., van der Ben, C. A., Roelen, C. A. M., Heymans, M. W., van der Klink, J. J. L., & Bültmann, U. (2016). Beyond return to work from sickness absence due to mental disorders: 5-year longitudinal study of employment status among production workers. *The European Journal of Public Health*, ckw178. https://doi.org/10.1093/eurpub/ckw178

Norrie, J., Heitger, M., Leathem, J., Anderson, T., Jones, R., & Flett, R. (2010). Mild traumatic brain injury and fatigue: A prospective longitudinal study. *Brain Injury*, *24*(13–14), 1528–1538. https://doi.org/10.3109/02699052.2010.531687

O'Keeffe, K., Hodder, S., & Lloyd, A. (2020). A comparison of methods used for inducing mental fatigue in performance research: Individualised, dual-task and short duration cognitive tests are most effective. *Ergonomics*, *63*(1), 1–12. https://doi.org/10.1080/00140139.2019.1687940

Otto, T., × Fred R H Zijlstra, & Goebel, R. (2018). Feeling the force: Changes in a left-lateralized network of brain areas under simulated workday conditions are reflected in subjective mental effort investment. *PLoS One*, *13*(6). Publicly Available Content Database. https://doi.org/10.1371/journal.pone.0198204

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Palacios, N., Molsberry, S., Fitzgerald, K. C., & Komaroff, A. L. (2023). Different risk factors distinguish myalgic encephalomyelitis/chronic fatigue syndrome from severe fatigue. *Scientific Reports*, *13*(1), Article 1. https://doi.org/10.1038/s41598-023-29329-x

Parmentier, F. B. R. (2014). The cognitive determinants of behavioral distraction by deviant auditory stimuli: A review. *Psychological Research*, *78*(3), 321–338. https://doi.org/10.1007/s00426-013-0534-4

Patel, A., Berdunov, V., Quayyum, Z., King, D., Knapp, M., & Wittenberg, R. (2020). Estimated societal costs of stroke in the UK based on a discrete event simulation. *Age and Ageing*, *49*(2), 270–276. https://doi.org/10.1093/ageing/afz162

Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiology & Behavior*, *93*(1–2), 369–378. https://doi.org/10.1016/j.physbeh.2007.09.016

Peelle, J. E. (2018). Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior. *Ear and Hearing*, *39*(2), 204–214. https://doi.org/10.1097/AUD.0000000000000494

Petruo, V. A., Mückschel, M., & Beste, C. (2018). On the role of the prefrontal cortex in fatigue effects on cognitive flexibility—A system neurophysiological approach. *Scientific Reports*, *8*(1), 6395. https://doi.org/10.1038/s41598-018-24834-w

Pezzullo, L., Streatfeild, J., Simkiss, P., & Shickle, D. (2018). The economic impact of sight loss and blindness in the UK adult population. *BMC Health Services Research*, *18*, 63. https://doi.org/10.1186/s12913-018-2836-0

Porter, S., Lexén, A., Johanson, S., & Bejerholm, U. (2018). Critical factors for the return-to-work process among people with affective disorders: Voices from two vocational approaches. *Work*, *60*(2), 221–234. https://doi.org/10.3233/WOR-182737

Proost, M., Habay, J., De Wachter, J., De Pauw, K., Rattray, B., Meeusen, R., Roelands, B., & Van Cutsem, J. (2022). How to Tackle Mental Fatigue: A Systematic Review of Potential Countermeasures and Their Underlying Mechanisms. *Sports Medicine*, *52*(9), 2129–2158. https://doi.org/10.1007/s40279-022-01678-z

Public Health England. (2018). *Briefing document: First incidence of stroke*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/678444/Stroke_incidence_briefing_document_2018.pdf

Qin, H., Zhou, X., Ou, X., Liu, Y., & Xue, C. (2021). Detection of mental fatigue state using heart rate variability and eye metrics during simulated flight. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *31*(6), 637–651. https://doi.org/10.1002/hfm.20927

Radford, K. A., Craven, K., McLellan, V., Sach, T. H., Brindle, R., Holloway, I., Hartley, S., Bowen, A., O'Connor, R., Stevens, J., Philips, J., Walker, M., Holmes, J., McKevitt, C., Murray, J., Watkins, C., Powers, K., Shone, A., & Farrin, A. (2020). An individually randomised controlled multi-centre pragmatic trial with embedded economic and process evaluations of early vocational rehabilitation compared with usual care for stroke survivors: Study protocol for the RETurn to work After stroKE (RETAKE) trial. *Trials*, *21*(1), 1010. https://doi.org/10.1186/s13063-020-04883-1

Ramage, A. E., Tate, D. F., New, A. B., Lewis, J. D., & Robin, D. A. (2019). Effort and Fatigue-Related Functional Connectivity in Mild Traumatic Brain Injury. *Frontiers in Neurology*, *0*. https://doi.org/10.3389/fneur.2018.01165

Reteig, L. C., van den Brink, R. L., Prinssen, S., Cohen, M. X., & Slagter, H. A. (2019). Sustaining attention for a prolonged period of time increases temporal variability in cortical responses. *Cortex*, *117*, 16–32. https://doi.org/10.1016/j.cortex.2019.02.016

Ricci, J. A., Chee, E., Lorandeau, A. L., & Berger, J. (2007). Fatigue in the U.S. Workforce: Prevalence and Implications for Lost Productive Work Time. *Journal of Occupational and Environmental Medicine*, *49*(1), 1–10. https://doi.org/10.1097/01.jom.0000249782.60321.2a

Riemersma, J. B. J., Biesta, P. W., & Wildervanck, C. (1978). *Fatigue and Stress Due to Prolonged Driving and Changing Task Demands*. 770134. https://doi.org/10.4271/770134

Riese, H. (1999). Mental Fatigue after Very Severe Closed Head Injury: Sustained Performance, Mental Effort, and Distress at Two Levels of Workload in a Driving Simulator. *Neuropsychological Rehabilitation*, *9*(2), 189–205. https://doi.org/10.1080/713755600

Rink, C., & Khanna, S. (2011). Significance of Brain Tissue Oxygenation and the Arachidonic Acid Cascade in Stroke. *Antioxidants & Redox Signaling*, *14*(10), 1889–1903. https://doi.org/10.1089/ars.2010.3474

Rönnbäck, L., & Johansson, B. (2022). Long-Lasting Pathological Mental Fatigue After Brain Injury–A Dysfunction in Glutamate Neurotransmission? *Frontiers in Behavioral Neuroscience*. Publicly Available Content Database. https://doi.org/10.3389/fnbeh.2021.791984

Rowe, F. J., Hepworth, L. R., Howard, C., Hanna, K. L., Cheyne, C. P., & Currie, J. (2019). High incidence and prevalence of visual problems after acute stroke: An epidemiology study with implications for service delivery. *PLoS ONE*, *14*(3), e0213035. https://doi.org/10.1371/journal.pone.0213035

Rowe, F. J., Hepworth, L. R., Howard, C., Hanna, K. L., & Currie, J. (2022). *Impact of Visual impairment following Stroke (IVIS study): A prospective clinical profile of central and peripheral visual deficits, eye movement abnormalities and visual perceptual deficits.*

Saar, K., Tolvanen, A., Poutiainen, E., & Aro, T. (2023). Returning to work after stroke: Associations with cognitive performance, motivation, perceived working ability and barriers. *Journal of Rehabilitation Medicine*, *55*, 2576. https://doi.org/10.2340/jrm.v55.2576

Sabes-Figuera, R., McCrone, P., Hurley, M., King, M., Donaldson, A. N., & Ridsdale, L. (2010). The hidden cost of chronic fatigue to patients and their families. *BMC Health Services Research*, *10*(1), 56. https://doi.org/10.1186/1472-6963-10-56

Saito, N., Yamano, E., Ishii, A., Tanaka, M., Nakamura, J., & Watanabe, Y. (2018). Involvement of the olfactory system in the induction of anti-fatigue effects by odorants. *PLoS One*, *13*(3). Publicly Available Content Database. https://doi.org/10.1371/journal.pone.0195263

Scholey, A. B., Benson, S., Sela-Venter, S., Mackus, M., & Moss, M. C. (2020). Oxygen Administration and Acute Human Cognitive Enhancement: Higher Cognitive

Demand Leads to a More Rapid Decay of Transient Hyperoxia. *Journal of Cognitive Enhancement*, *4*(1), 94–99. https://doi.org/10.1007/s41465-019-00145-4

Segerstrom, S. C., & Nes, L. S. (2007). Heart rate variability reflects self-regulatory strength, effort, and fatigue. *Psychological Science*, *18*(3), 275–281. https://doi.org/10.1111/j.1467-9280.2007.01888.x

Sengupta, A., Tiwari, A., & Routray, A. (2017). Analysis of cognitive fatigue using EEG parameters. *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2554–2557. https://doi.org/10.1109/EMBC.2017.8037378

Shan, Z. Y., Finegan, K., Bhuta, S., Ireland, T., Staines, D. R., Marshall-Gradisnik, S. M., & Barnden, L. R. (2018). Brain function characteristics of chronic fatigue syndrome: A task fMRI study. *NeuroImage: Clinical*, *19*, 279–286. https://doi.org/10.1016/j.nicl.2018.04.025

Shelton, B., Nesbitt, K., Thorpe, A., & Eidels, A. (2021). Gauging the utility of ambient displays by measuring cognitive load. *Cognition, Technology & Work*, *23*(3), 459–480. https://doi.org/10.1007/s10111-020-00639-8

Sinclair, E., Radford, K., Grant, M., & Terry, J. (2014). Developing stroke-specific vocational rehabilitation: A soft systems analysis of current service provision. *Disability and Rehabilitation*, *36*(5), 409–417. https://doi.org/10.3109/09638288.2013.793410

Slimani, M., Znazen, H., Bragazzi, N., Zguira, M., & Tod, D. (2018). The Effect of Mental Fatigue on Cognitive and Aerobic Performance in Adolescent Active Endurance Athletes: Insights from a Randomized Counterbalanced, Cross-Over Trial. *Journal of Clinical Medicine*, *7*(12), 510. https://doi.org/10.3390/jcm7120510

Smith, M. R., Chai, R., Nguyen, H. T., Marcora, S. M., & Coutts, A. J. (2019). Comparing the Effects of Three Cognitive Tasks on Indicators of Mental Fatigue. *The Journal of Psychology*, *153*(8), 759–783. https://doi.org/10.1080/00223980.2019.1611530

Souchet, A. D., Philippe, S., Lourdeaux, D., & Leroy, L. (2022). Measuring Visual Fatigue and Cognitive Load via Eye Tracking while Learning with Virtual Reality Head-Mounted Displays: A Review. *International Journal of Human–Computer Interaction*, *38*(9), 801–824. https://doi.org/10.1080/10447318.2021.1976509

Sun, L., Guo, Z., Yuan, X., Wang, X., Su, C., Jiang, J., & Li, X. (2022). An Investigation of the Effects of Brain Fatigue on the Sustained Attention of Intelligent Coal Mine VDT Operators. *International Journal of Environmental Research and Public Health*, *19*(17), 11034. https://doi.org/10.3390/ijerph191711034

Tanaka, M. (2015). Effects of Mental Fatigue on Brain Activity and Cognitive Performance: A Magnetoencephalography Study. *Anatomy & Physiology*, *s4*. https://doi.org/10.4172/2161-0940.S4-002

Tanaka, M., Mizuno, K., Yamaguti, K., Kuratsune, H., Fujii, A., Baba, H., Matsuda, K., Nishimae, A., Takesaka, T., & Watanabe, Y. (2011). Autonomic nervous alterations associated with daily level of fatigue. *Behavioral and Brain Functions*, *7*, 46. Publicly Available Content Database. https://doi.org/10.1186/1744-9081-7-46

Tanaka, M., Yamada, H., Nakamura, T., Ishii, A., & Watanabe, Y. (2013). Fatigue-recovering effect of a house designed with open space. *Explore (New York, N.Y.)*, *9*(2), 82-86. https://doi.org/10.1016/j.explore.2012.12.006

Tate, R. L., Godbee, K., & Sigmundsdottir, L. (2013). A systematic review of assessment tools for adults used in traumatic brain injury research and their relationship to the ICF. *NeuroRehabilitation*, *32*(4), 729–750. https://doi.org/10.3233/NRE-130898

Tay, S. S., Visperas, C. A., Tan, M. M. J., Chew, T. L. T., & Koh, X. H. (2023). Factors Predicting Return to Work After Inpatient Stroke Rehabilitation: A Retrospective Follow-up Study. *Archives of Rehabilitation Research and Clinical Translation*, *5*(1), 100253. https://doi.org/10.1016/j.arrct.2022.100253

Terentjeviene, A., Maciuleviciene, E., Vadopalas, K., Mickeviciene, D., Karanauskiene, D., Valanciene, D., Solianik, R., Emeljanovas, A., Kamandulis, S., & Skurvydas, A. (2018). Prefrontal Cortex Activity Predicts Mental Fatigue in Young and Elderly Men During a 2 h "Go/NoGo" Task. *Frontiers in Neuroscience*. Publicly Available Content Database. https://doi.org/10.3389/fnins.2018.00620

Theadom, A., Rowland, V., Levack, W., Starkey, N., Wilkinson-Meyers, L., & McPherson, K. (2016). Exploring the experience of sleep and fatigue in male and female adults over the 2 years following traumatic brain injury: A qualitative descriptive study. *BMJ Open*, *6*(4), e010453. https://doi.org/10.1136/bmjopen-2015-010453

Thiffault, P., & Bergeron, J. (2003). Fatigue and individual differences in monotonous simulated driving. *Personality and Individual Differences*, *34*(1), 159–176. https://doi.org/10.1016/S0191-8869(02)00119-8

Thill, S., Houssemand, C., & Pignault, A. (2020). Effects of meaning in life and of work on health in unemployment. *Health Psychology Open*, *7*(2), 2055102920967258. https://doi.org/10.1177/2055102920967258

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

Tomar, S., Sharma, A., Jain, A., Sinha, V. D., & Gupta, I. D. (2018). Study of Fatigue and Associated Factors in Traumatic Brain Injury and Its Correlation with Insomnia and

Depression. *Asian Journal of Neurosurgery*, *13*(4), 1061–1065. https://doi.org/10.4103/ajns.AJNS_89_17

Tran, Y., Wijesuriya, N., Tarvainen, M., Karjalainen, P., & Craig, A. (2009). The Relationship Between Spectral Changes in Heart Rate Variability and Fatigue. *Journal of Psychophysiology*, *23*(3), 143–151. https://doi.org/10.1027/0269-8803.23.3.143

Truszczynski, O., Wojtkowiak, M., Biernacki, M., & Kowalczuk, K. (2009). The Effect of Hypoxia on the Critical Flicker Fusion Threshold in Pilots. *International Journal of Occupational Medicine and Environmental Health*, *22*(1), 13–18. Publicly Available Content Database.

Tsao, J., Evans, S., Seidman, L., Lung, Zeltzer, L., & Naliboff, B. (2013). Heart rate variability as a biomarker for autonomic nervous system response differences between children with chronic pain and healthy control children. *Journal of Pain Research*, 449. https://doi.org/10.2147/JPR.S43849

Turner, R. (2013). The handbook of operator fatigue. *Ergonomics*, *56*(9), 1486–1486. https://doi.org/10.1080/00140139.2013.819674

Van Cutsem, J., De Pauw, K., Vandervaeren, C., Marcora, S., Meeusen, R., & Roelands, B. (2019). Mental fatigue impairs visuomotor response time in badminton players and controls. *Psychology of Sport and Exercise*, *45*, N.PAG. https://doi.org/10.1016/j.psychsport.2019.101579

van der Schaaf, M. E., Roelofs, K., de Lange, F. P., Geurts, D. E. M., van der Meer, J. W. M., Knoop, H., & Toni, I. (2018). Fatigue Is Associated With Altered Monitoring and Preparation of Physical Effort in Patients With Chronic Fatigue Syndrome. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(4), 392–404. https://doi.org/10.1016/j.bpsc.2018.01.015

Völker, I., Kirchner, C., & Bock, O. L. (2016). On the relationship between subjective and objective measures of fatigue. *Ergonomics*, *59*(9), 1259–1263. https://doi.org/10.1080/00140139.2015.1110622

Vrijkotte, S., Meeusen, R., Vandervaeren, C., Buyse, L., Cutsem, J., Pattyn, N., & Roelands, B. (2018). Mental Fatigue and Physical and Cognitive Performance During a 2-Bout Exercise Test. *International Journal of Sports Physiology and Performance*, *13*(4), 510-516.

Walker, L. A. S., Lindsay-Brown, A. P., & Berard, J. A. (2019). Cognitive Fatigability Interventions in Neurological Conditions: A Systematic Review. *Neurology and Therapy*, *8*(2), 251–271. https://doi.org/10.1007/s40120-019-00158-3

Wang, F., Wang, H., & Fu, R. (2018). Real-Time ECG-Based Detection of Fatigue Driving Using Sample Entropy. *Entropy*, *20*(3), 196. Publicly Available Content Database. https://doi.org/10.3390/e20030196

Wascher, E., Heppner, H., Kobald, S. O., Arnau, S., Getzmann, S., & Möckel, T. (2016). Age-Sensitive Effects of Enduring Work with Alternating Cognitive and Physical Load. A Study Applying Mobile EEG in a Real Life Working Scenario. *Frontiers in Human Neuroscience*. Publicly Available Content Database. https://doi.org/10.3389/fnhum.2015.00711

Watts, M. E., Pocock, R., & Claudianos, C. (2018). Brain Energy and Oxygen Metabolism: Emerging Role in Normal Function and Disease. *Frontiers in Molecular Neuroscience*, *11*. https://www.frontiersin.org/articles/10.3389/fnmol.2018.00216

Williamson, A. M., Feyer, A.-M., & Friswell, R. (1996). The impact of work practices on fatigue in long distance truck drivers. *Accident Analysis & Prevention*, *28*(6), 709–719. https://doi.org/10.1016/S0001-4575(96)00044-9

Wilson, N., Guragain, B., Verma, A., Archer, L., & Tavakolian, K. (2020). Blending Human and Machine: Feasibility of Measuring Fatigue Through the Aviation Headset. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *62*(4), 553–564. https://doi.org/10.1177/0018720819849783

Xiao, Y., Wu, J., Li, J., Tang, W., Ma, F., Sun, C., Yang, Y., Zhan, W., Wang, L., Yan, H., Xu, F., & Chen, S. (2019). The Neuro Patterns Prior to Error Responses in Long-Lasting Working Memory Task: An Event-Related Potential Study. *Frontiers in Behavioral Neuroscience*, *13*, 277. https://doi.org/10.3389/fnbeh.2019.00277

Yoon, J.-H., Park, N.-H., Kang, Y.-E., Ahn, Y.-C., Lee, E.-J., & Son, C.-G. (2023). The demographic features of fatigue in the general population worldwide: A systematic review and meta-analysis. *Frontiers in Public Health*, *11*, 1192121. https://doi.org/10.3389/fpubh.2023.1192121

Yu, S., Mückschel, M., & Beste, C. (2022). Superior frontal regions reflect the dynamics of task engagement and theta band-related control processes in time-on task effects. *Scientific Reports*.

Zarei, K., Poolman, P., Garvin, M. K. & Kardon, R. (2017). eyeFusion: Mobile, Objective Quantification of Critical Flicker Fusion Thresholds. *Investigative Ophthalmology & Visual Science, 58*(3129). https://iovs.arvojournals.org/article.aspx?articleid=2638026

Zargari, M. R., Pascal, M., Øyvind, O., Nicolas, V. & Afshin, S. (2018). Eye movement characteristics reflected fatigue development in both young and elderly individuals. *Scientific Reports (Nature Publisher Group)*, *8*(1). Publicly Available Content Database. https://doi.org/10.1038/s41598-018-31577-1

Zhang, C., & Yu, X. (2010). Estimating mental fatigue based on electroencephalogram and heart rate variability. *Polish Journal of Medical Physics and Engineering*, *16*(2), 67. Publicly Available Content Database. https://doi.org/10.2478/v10013-010-0007-7

Zhang, C., Zheng, C.-X., & Yu, X.-L. (2008). Automatic recognition of cognitive fatigue from physiological indices by using wavelet packet transform and kernel learning algorithms. *Expert Systems with Applications*, *36*(3), 4664–4671. https://doi.org/10.1016/j.eswa.2008.06.022

Zhang, H., Wang, J., Geng, X., Li, C., & Wang, S. (2021). Objective Assessments of Mental Fatigue During a Continuous Long-Term Stress Condition. *Frontiers in Human Neuroscience*, *15*, 733426. https://doi.org/10.3389/fnhum.2021.733426

Zhang, Q., Sun, M. A., Sun, Q., Mei, H., Rao, H., & Liu, J. (2023). Mental Fatigue Is Associated with Subjective Cognitive Decline among Older Adults. *Brain Sciences*, *13*(3), 376. https://doi.org/10.3390/brainsci13030376

Zhang, S., Hu, S., Chao, H. H., Luo, X., Farr, O. M., & Li, C. R. (2012). Cerebral correlates of skin conductance responses in a cognitive task. *NeuroImage*, *62*(3), 1489–1498. https://doi.org/10.1016/j.neuroimage.2012.05.036