Dickens, Alicia, Champion, Andrew J and Schenke, Kimberley C ORCID logoORCID: https://orcid.org/0000-0002-1184-4802 (2024) Fatigue management: a systematic review of objective measurement techniques for cognitive fatigue. Journal of Clinical and Experimental Neuropsychology, 46 (8). pp. 776-793. doi:10.1080/13803395.2024.2415070

Official URL: http://dx.doi.org/10.1080/13803395.2024.2415070
DOI: http://dx.doi.org/10.1080/13803395.2024.2415070
EPrint URI: https://eprints.glos.ac.uk/id/eprint/14462

Fatigue management: A Systematic Review of Objective Measurement Techniques for Cognitive Fatigue

Alicia Dickens[a], Andrew J. Champion[a] and Kimberley C. Schenke[a]

[a]University of Gloucestershire

Corresponding author: Dr Kimberley Schenke, kschenke1@glos.ac.uk

University of Gloucestershire, Swindon Road, Cheltenham, GL50 4AZ

Abstract

Background

Cognitive fatigue is a complex psychobiological state whereby task performance cannot be maintained. Return-to-work protocols typically rely on self-report measures, therefore the current systematic review aimed to identify 'real-time' measures of objective cognitive fatigue to inform return-to-work protocols.

Methods

Studies were included if participants were at least 18 years old, assessed 'real-time' objective cognitive fatigue that could be used outside of the lab (neuroimaging measures were, therefore, excluded), used an induction task that was separate to the measurement, were adequately powered, compared objective cognitive fatigue at baseline and post-induction, and included a cognitive fatigue induction task that was at least 30 minutes long.

Nine electronic databases were searched until December 31, 2022 (MEDLINE; PsychArticle; PubMED, ProQuest; ProQuest for grey literature; Google Scholar; The Cochrane Library; The

Health Technology Assessment Database; and Web of Science), with alerts set up on Google Scholar to notify of new relevant research since this date (reviewed until December 2023).
The checklist for quasi-experimental studies (Joanna Briggs Institute, 2014) was used to assess risk of bias. Whilst a meta-analysis was planned, the data were unsuitable so only a narrative synthesis was conducted.

Results

Fifty-seven studies were included, which were conducted within a variety of settings including naturalistic work scenarios, driving, aviation, and artificial computer-based tasks.
Whilst the review found a range of potential measurements, there were inconsistencies in findings across studies highlighting the need for more research into the reliable measurement of objective cognitive fatigue in natural settings.

Discussion

The findings suggest that eye- and body-related measures may be sensitive measures of objective cognitive fatigue. However, comparisons across measurement types should be cautiously interpreted because eye-related and cognitive measures were far more common. The review highlighted the need for more consistent and transparent reporting across the field to advance our understanding of cognitive fatigue.

Keywords: Cognitive Fatigue, Mental Fatigue, Return To Work, Return To Meaningful Occupation, Return To Activity

Introduction

Cognitive fatigue is commonly described as a psychobiological state in which task performance cannot be maintained due to a decrease in available cognitive resources (Borragán et al., 2017). This can be attributed to an inability to allocate resources effectively (Proost et al., 2022) and/or reduced motivation to redistribute efforts (Herlambang et al., 2021), enabling categorisation into 'objective' (biological) fatigue (cognitive resources are low, but motivation is unaffected) and 'subjective' (psychological) fatigue (cognitive resources are adequate, but motivation is low; Benoit et al., 2017).

Cognitive fatigue is a common experience (Walker et al., 2019) that is often 'recovered' in a neurotypical person following sufficient rest (Mizuno et al., 2011). However, without adequate rest cognitive fatigue can hamper task completion (Walker et al., 2019) and work pattern maintenance (Hiploylee et al., 2017; Losoi et al., 2016) due to reduced attention and alertness (Muñoz-de-Escalona et al., 2020). Cognitive fatigue increases accident/injury risk (Proost et al., 2022), and can be deemed pathological when it prevents the completion of daily tasks (Rönnbäck & Johansson, 2022).

Cognitive fatigue accompanies many illnesses and long-term conditions (Dantzer et al., 2014), and is commonly associated with brain injury and stroke (Johansson & Rönnbäck, 2012; Norrie et al., 2010). The potential impact on activities of daily living and workplace safety makes accurate measurement of cognitive fatigue important (Guo et al., 2016), especially in those with identified difficulties allocating (reduced) cognitive resources such as brain injury or stroke populations, who may also have less insight into their cognitive fatigue (Ramage et al., 2019), or be less able to answer self-report measures accurately, exacerbating potential risks (Ramage et al., 2019; Ricci et al., 2007).

Whilst clinicians involved in return-to-work protocols are often reliant on self-report measures (Benoit et al., 2017; NICE, 2021), there can be discrepancies between self-report and objective ability to perform tasks (Ricci et al., 2007). For example, responses to self-report measures of fatigue typically suggest that respondents are less fatigued than objective measures would indicate. This discrepancy has the potential to cause distress and/or danger for individuals who may lack insight into the extent of their cognitive fatigue and its impact upon performance (Ricci et al., 2007). Additionally, self-report responses within return-to-work protocols may be called into question if there are incentives to remain out of employment. Moreover, there is often conflation between self-report measures of physical and cognitive fatigue. Physical fatigue, as with cognitive fatigue, is often assessed through self-report questionnaires in return-to-work contexts. Many of these measures include items addressing both physical and cognitive fatigue, often combining them into an overall total while acknowledging a multidimensional understanding of fatigue. This overlap emphasises the importance of establishing cognitive fatigue as a standalone indicator of the need for reasonable adjustments in the workplace. Fatigue management recommendations for return to work should, therefore, incorporate objective monitoring of cognitive fatigue (Dawson et al., 2011; Hopstaken, 2016).

This systematic review aims to identify 'real-time' measures of objective cognitive fatigue that are applicable in 'real-life' settings, focusing on measures that could inform return-to-work protocols following brain injury or stroke. This would enable individuals to monitor their own performance (Guo et al., 2016), increasing agency over their return to activity. In contexts such as operating machinery, driving and healthcare-related decision-making (Dawson & Fletcher, 2001; Guo et al., 2016), this takes on additional importance. There are two key objectives for the review:

1. Identify 'real-time' objective measures of cognitive fatigue that do not require a laboratory setting or expensive equipment
2. Assess the reliability and validity of these measures

Systematic Review Methods

This systematic review is registered on the PROSPERO database (CRD42022323292) and was conducted according to the Cochrane Handbook for Systematic Reviews (Higgins et al., 2019) and the Preferred Reporting for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2010).

Information Sources and Search Strategy

Nine electronic databases were searched until December 31, 2022 (MEDLINE; PsychArticle; PubMED, ProQuest; ProQuest for grey literature; Google Scholar; The Cochrane Library; The Health Technology Assessment Database; and Web of Science), with alerts set up on Google Scholar to notify of new relevant research since this date (reviewed until December 2023). The reference sections of eligible articles were hand-searched for further relevant records. To maximise search results, the keywords used for all database searches were kept broad: "Mental fatigue" OR "cognitive fatigue" OR "mental exertion" OR "cognitive exertion" OR "cognitive load" AND "objective" AND "measure".

Eligibility Criteria and Study Selection

The PICOTS framework (e.g., Riva et al., 2012) specified the following, a priori, criteria:

- **Population:** Participants at least 18 years of age.
- **Intervention:** Studies assessing 'real-time' objective cognitive fatigue using an induction task that was separate to the measurement to avoid the confound of the measure itself cognitively fatiguing participants. As one objective is to assess the

measures' reliability/validity, a further eligibility criterion was that studies should be powered to at least .80 power, so case studies and pilot studies were excluded.

- **Comparator:** Comparison between objective cognitive fatigue level at baseline and post-induction.
- **Outcome:** The primary outcome should be the analysis of difference between objective measures of cognitive fatigue pre- and post- fatigue induction. A secondary outcome should be subjective cognitive fatigue measurement pre- and post- fatigue induction. Neuroimaging measures were excluded due to their limited applicability in 'real-world' settings.
- **Time:** The duration of the cognitive fatigue induction task had to be a minimum of 30 minutes (Borragán et al., 2017; Holgado et al., 2020; Tanaka, 2015). If the duration was unclear or not reported then the study was excluded.
- **Setting:** The measures should be practical within a naturalistic setting.

The initial literature search results were compiled in Microsoft Excel and duplicates removed. Two authors (KS, AD) independently screened the titles and abstracts for the following criteria (disagreements resolved by a third author; AC):

1. The publication was written in English
2. Objective measures of cognitive fatigue in humans were used
3. Original primary data was collected (review/theoretical articles and analyses of secondary data were excluded)

The same two authors then independently screened the full text of the eligible articles identified against the following criteria (disagreements resolved by the third author):

1. There was primary data for ‘real-time’ objective cognitive fatigue in humans (i.e., objective cognitive fatigue was induced during the study).
2. The study was a full empirical study, and not a case study, pilot study or proof of concept study.
3. The task used to induce cognitive fatigue was independent of the measurement of objective cognitive fatigue.
4. If objective cognitive fatigue was measured using neuroimaging there was at least one other objective measure utilised.
5. The induction task was at least 30 minutes long.

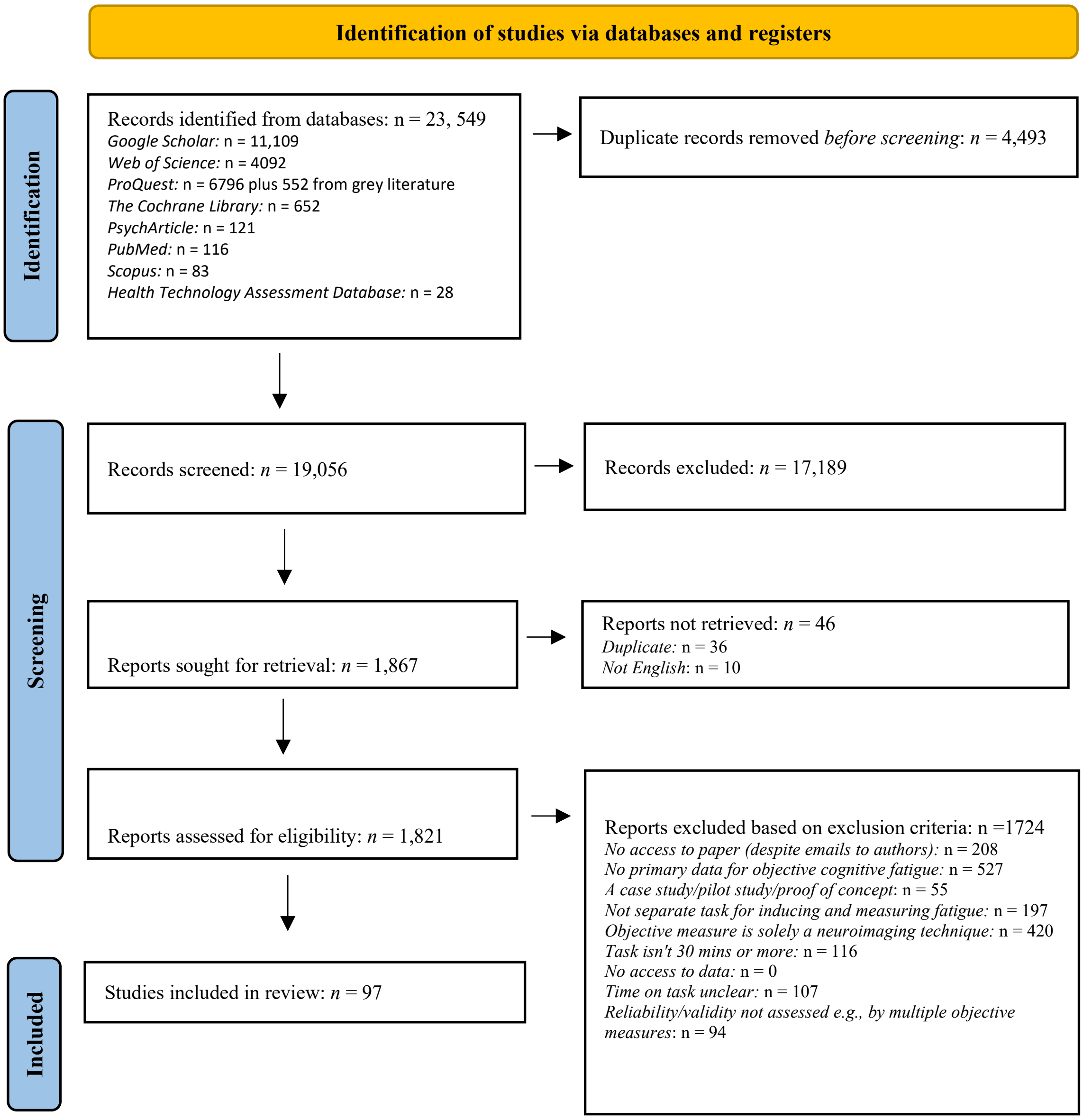
Given the aim of identifying a *reliable* objective measure of cognitive fatigue, the following additional criterion was applied as the initial search and primary screening indicated inconsistencies of findings:

6. Studies included more than one objective measure of cognitive fatigue, enabling an assessment of reliability.

The PRISMA flow diagram (based on Page et al., 2021) in Figure 1 gives the results of the screenings: 97 studies were included.

Figure 1

PRISMA flow diagram detailing the results of the screening protocols



The references of all eligible articles were hand searched for further eligible articles, adding 23 articles. The references of these 23 articles were hand searched, adding 2 more. No further studies were added from hand searching these 2 articles. A total of 122 journal articles were eligible for extraction.

Data Extraction

Two authors (AD, KS) extracted the following information from each eligible study: method and duration of cognitive fatigue induction; objective measures used to assess cognitive fatigue; descriptive and inferential statistics comparing cognitive fatigue pre- and post-induction; study title; authors; year published; number and characteristics of participants; study design; aims and hypotheses; conclusion; reported strengths and limitations.

Two articles were excluded during extraction for being duplicates, 44 for not meeting the eligibility requirements/not having a baseline comparison/providing incomplete data/analysis, and one for including participants under 18. Thus, risk of bias was assessed for 79 studies within 75 articles.

Risk of Bias

The checklist for quasi-experimental studies (Joanna Briggs Institute, 2014) was selected as an appropriate risk of bias tool. Two of its nine criteria ('were outcomes measured in a reliable way?' and 'was appropriate statistical analysis used?') were given higher weighting due to the review's aims. However, few studies reported a priori power analyses, making the latter question difficult to assess. To avoid excluding the majority of studies, rudimentary 'a priori' power analyses were conducted using G Power (Faul et al., 2007) to provide an indication as to whether studies were adequately powered. Whilst this approach may miss some nuances within the designs, it was adopted given the lack of information within most of the articles around power, design and the provided statistics. As the literature typically suggests a large effect size (Borragán et al., 2017; O'Keeffe et al., 2020), these power analyses were calculated based on a large effect size (.80) for .80 power with an alpha level of .05 (see Table 1). For the power analyses that required estimation of the correlation between repeated measures variables a moderate positive correlation estimate of .60 was used. Studies were rated as not meeting the ‘appropriate statistical analysis’ criterion if they were considered underpowered.

Table 1

The minimum sample estimates for the most commonly used statistical test based on simulated a priori power analyses within GPower

Statistical test	**Minimum sample size suggested**
t test	15
one way ANOVA (3 levels)	10
2 x 2 ANOVA	12
2 x 2 x 2 ANOVA	16

Twenty-three studies were judged to be underpowered and were removed from the narrative synthesis. Thus, 57 studies were included within the narrative synthesis from 52 journal articles.

Data synthesis

The narrative synthesis encompassed two aspects; sample characteristics and study characteristics (with induction task and measures considered separately).

Sample characteristics

Fifteen studies did not specify their population. Sixteen recruited students (three required participants to be drivers, one recruited participants with dysphoria and one recruited only males), two recruited pilots/aviators, two recruited truck drivers, one recruited amateur boxers who were non-smokers, one recruited experienced table tennis players, one recruited badminton players, and one recruited construction equipment operators. Three studies only specified that participants had driving experience; an additional three studies required the drivers to be male, with one of these also stipulating right-handedness. One study required participants to be male non-smokers; two only recruited non-smokers (one of which required completion of at least one sudoku puzzle a month). Four studies included participants with a health condition (head injury in two studies, Multiple Sclerosis and/or stroke in two studies), and one required participants to be sleep-deprived. Two studies required a group of ‘young’ participants and a group of ‘older’ participants, and one study required participants to be right-handed adults.

Nine studies used a male-only population (of these, one recruited graduate students, one recruited non-smokers and one recruited right-handed drivers). Eight studies did not report participants’ sex or gender. Across the studies the mean number of males was 16.0 (*SD* = 9.3) and the mean number of females was 13.2 (*SD* = 14.3). The average age of all participants was 27.1 years old with an SD of 6.8, ranging between 18 and 91 years old.

The average sample size across the studies was 27.3 (*SD* = 15.8), ranging from 10 to 82. Only four studies discussed prospective power analyses to justify their sample size. Table 2 summarises sample characteristics.

Table 2

A summary of the sample characteristics ordered by the year of publication

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)	Long distance truck drivers	37	N/A	37.4
Williamson et al. (1996)	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 (2009)	Drivers	24	M = 12, F = 12	N/A (24-30)
Pattyn et al. (2008)	Students	21	N/A	N/A
Tanaka et al. (2009)	Males	20	M = 20, F = 0	20.6
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)	Non-smokers who play sudoku	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 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)	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 right-handed drivers	35	M = 35, F = 0	22
Guo et al. (2018)	University students who drive	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)	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 & elderly groups	38	M = 18, F = 20	40.5
Zhang et al. (2018)	Male graduate students	14	M = 14, F = 0	23 (21-26)
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 Junior (2020)	N/A	20	M = 10, F = 10	25.6 (18-35)
Herlambang et al. (2021)	University students	28	M = 17, F = 11	24.57
Wilson et al. (2020)	Sleep deprived	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)	Non-smoking amateur boxers	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
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

Studies used various contexts. Four investigated naturalistic scenarios in participants' work settings, 20 investigated driving, 6 investigated aviation (real or simulated), and 27 involved artificial computer-based tasks. The earliest study was published in 1978 and the most recent in 2022 (median year = 2018). Only one study mentioned a likely paper-based task, which was a reading comprehension task. All other tasks were explicitly computerised or involved simulations, indicating a strong preference for digitised methods in these studies. Most tasks were heavily based on vision, though there were a few tasks that were auditory in nature.

Definitions of cognitive fatigue

Although not a specific objective of the systematic review itself, the definitions of cognitive fatigue within the eligible studies were extracted. Of the 53 articles included, only 25 explicitly defined cognitive fatigue. Six of these defined a specific task-based aspect of fatigue, such as ‘driver’ (four studies) or ‘pilot’ fatigue (two studies), tending to focus on reductions in alertness and attention, and increased response times. Various terms relevant to cognitive fatigue were used somewhat interchangeably within the literature including ‘mental effort’ and ‘cognitive load’.

Some definitions alluded to an unwillingness to continue with a task (i.e., the motivational aspect of cognitive fatigue) and focused on the subjective experience of exhaustion accompanying a reduction in task performance. In these definitions, the focus was on individuals being aware that they had a decreased capacity for the task. Other definitions focused more on the objective, observable manifestations of cognitive fatigue such as the reduction in the efficiency of mental processes (e.g., reduced attention, increased distractibility and impaired information processing). Others still focused on psychophysiological changes within cognitive fatigue and the relationship between the available mental resources and arousal or motivation. Clearer distinctions between cognitive fatigue and visual fatigue would allow for a better understanding of their relative impacts upon eye-related measures (Souchet et al., 2022).

Cognitive Fatigue Induction Procedures

The cognitively fatiguing tasks ranged from 30 to 480 minutes (*mean* = 94.1, *SD* = 89.2, *median* = 60, *mode* = 60). There were four outliers (two studies had tasks for 240 minutes, two had tasks for 480 minutes) within the otherwise negatively skewed data. Excluding these outliers, most studies had a duration between 30 and 60 minutes (*mean* = 74.1, *SD* = 40.1, *median* = 60, *mode* = 49.4). While 26 studies did not include a post-task measurement, most of those that did report that the post-task assessment occurred immediately after the induction tasks. For the remaining studies, although they did not explicitly state the timing, in the absence of any mention of a delay it is assumed that the post-task measurement took place immediately.

Approximately half of the studies (29) incorporated a task specifically designed to induce fatigue (including the Stroop task [3], number-based tasks [4], n-back task [12], switch tasks [1], continuous performance tasks [2], the go no-go task [1], pattern recognition [1], reading comprehension [1], video games or social media [1], and the cognitive vigilance task [1] – note, two studies used several of these tasks) and then measured performance on another task. The other half of the studies had participants undertaking one task continuously whilst having at least one secondary task that either occurred concurrently or immediately after the fatigue induction (see Table 3). These 'prolonged' tasks involved driving (19), aviation (six) or workplace activities (three).

Table 3

A summary of the study characteristics, ordered by year of publication

	Induction				Differences from baseline measures	
Citation	**Context/Setting**	**Task**	**Time (mins)**	**Objective measure**	**Objective**	**Subjective**
Riemersma et al. (1978a)	Driving	Driving	480	Intermittent reporting task (report kms travelled), Colour changing task, HR	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	Driving	150	HR, episodic secondary task, critical flicker fusion freq., vigilance, unstable tracking	Yes (some measures)	Unclear
Williamson et al. (1996)	Driving	Driving	149	vigilance, critical flicker fusion, unstable tracking, HR	Yes (some measures)	Yes
Riese et al. (1999)	Driving simulator	Driving	50	Selective attention task, HR, BP, Priming and recognition task, Memory search task, Dot counting task, Peripheral stimuli	Yes (some measures)	Yes
Li et al. (2004)	Driving	Driving	180	Vigilance (RTs), HR Variability, critical flicker fusion freq.	Yes (some measures)	Unclear
LeDuc et al. (2005)	Aviation	Helicopter flight	72	Pupil size, pupil constriction, latency, constriction amplitude, saccadic velocity	Yes (some measures)	Yes
Liu & Wu (2009)	Driving Simulator	Driving	60	Arithmetic task performance, Distance estimation task	Yes (some measures)	Yes
Pattyn et al. (2008)	Laboratory	Cognitive vigilance	90	Respiratory inductive plethysmography (abdominal & ribcage excursions)	Yes (some measures)	Unclear
Tanaka et al. (2009)	Laboratory	n-back	30	Advanced Trail Making Task, HR	Yes (some measures)	N/A
Claros-Salinas et al. (2010)	Laboratory	Normal daily routine	60	Go no go task, Divided attention task, Simple reaction time task	Yes (some measures)	Yes
Massar et al. (2010)	Driving simulator	n-back	90	Driving performance, EEG, Auditory oddball task	Yes (some measures)	Yes
Zhang & Yu (2010)	Laboratory	Mental arithmetic task	120	EEG, HR	Yes	Yes
Mckinley et al. (2011)	Aviation	Unmanned aerial vehicle target acquisition	30	Psychomotor visual, landing, PERCLOS, Approx. Entropy for pupil position	Yes (some measures)	N/A
Tanaka et al. (2011)	Laboratory	n-back	30	Advanced trail making test, ECG	Yes (some measures)	Yes
Jagannath & Balasubramanian (2014)	Driving Simulator	Driving	60	Surface EMG, EEG, seat interface pressure, BP, HR, oxygen saturation level	Yes (some measures)	Unclear
Gergelyfi et al., (2015)	Laboratory	Sudoku	120	EEG, ECG-HR Variability, Pupil diameter, Eye blink, skin conductance response, working memory task, Eye movements	Yes (some measures)	Yes
Guo et al. (2015)	Workplace/daily life	AX-CPT	60	Go no go task, EEG	Yes (some measures)	Yes
Hopstaken et al. (2015)	Laboratory	n-back	60	Pupil diameter, stimulus-evoked pupil dilation	Yes	Yes
Körber et al. (2015)	Driving Simulator	Driving	42.5	Vigilance, blink frequency, blink duration, pupil diameter, PERCLOS	Yes (some measures)	N/A
Hopstaken et al. (2016)	Laboratory	n-back	90	Gaze position, Pupil diameter, EEG	Yes (some measures)	Yes
Völker et al. (2016)	Workplace	Clerical work	480	PVT, number of characters typed, typing errors, eye blinks, posture, observer rated facial features	Yes (some measures)	Yes

Zhao et al. (2017)	Driving Simulator	Driving	90	EEG, oddball task, ECG, blink frequency	Yes	Yes
Balasubramanian & Bhardwaj (2018)	Driving Simulator	Driving	120	EEG, ECG	Yes	N/A
Guo et al. (2018)	Driving Simulator	Driving	90	Go no go task, EEG	Yes	Yes
Jonasson et al. (2018)	Laboratory	Reading comprehension	30	Digit symbol coding, attentional blink test, computerized rest combining divided attention and working memory simultaneously, digit span, symbol search, HR, Stroop	Yes (some measures)	Unclear
Ma et al. (2018)	Driving Simulator	Driving	60	EEG, Eye measures	Yes (some measures)	Yes
Otto et al. (2018)	Aviation simulator	Helicopter training	240	n back, MRI	Yes	Yes
Saito et al. (2018)	Laboratory	n-back	40	Simple colour reaction task, Stroop, Accelerated plethysmography	No	Yes
Wang et al. (2018)	Driving simulator	Driving	240	ECG, EEG	Yes	Unclear
Zargari et al. (2018)	Laboratory	Pattern recognition	40	Blink duration, Blink frequency, saccade duration, Peak velocity of eye movement, Pupil dilation range	Yes	Yes
Zhang et al. (2018)	Laboratory	Switch task	120	EEG, HR Variability	Yes (some measures)	Yes
Herlambang et al. (2019)	Laboratory	Vowel counting in pseudowords	150	Eye diameter, gaze positions, saccades, blinks, ECG	Yes (some measures)	Yes (some measures)
Smith et al. (2019)	Laboratory	AX-CPT, Stroop & Psychomotor vigilance	45	EEG, HR Variability, 3 min PVT	Yes (some measures)	Yes
Van Cutsem et al. (2019)	Laboratory	Stroop	90	Visuomotor task, Flanker task, HR	Yes (some measures)	Yes
Ferreira Junior (2020)	Laboratory	Stroop	30	HR Variability, Reaction time task	Yes	Yes
Herlambang et al. (2021)	Laboratory	Sudoku	39	Pupil diameter, eyeblinks, gaze positions, saccades, HR Variability	Yes	Yes
Wilson et al. (2020)	Flight Simulator	Flying (take off, climb & cruise phase)	126	HR Variability, photoplethysmogram	Yes	N/A
Chen et al. (2021)	Laboratory	n-back	100	EEG, Suspension task 9some), Maths task (some)	Yes (some measures)	Yes
Fortes et al. (2021)	Sport/laborato ry	Coaching video/social media/video game	30	Stroop, Boxing decision-making	Yes (some measures)	Yes
Habay et al. (2021)	Laboratory	Stroop	60	Flanker task, EEG, ECG, Visuomotor performance	Yes (some measures)	Yes (some measures)
Hu & Lodewijks (2021a)	Driving Simulator	Simulated driving	60	Pupil size, Vigilance (SDT), blinks, eye fixation, eye saccades	Yes (some measures)	Yes
Hu & Lodewijks (2021b)	Driving	Simulated driving	60	Pupil size, Vigilance (SDT), blinks, eye fixation, eye saccades	Yes (some measures)	Yes
Magnuson et al. (2021)	Laboratory	n-back	60	EEG, go/no go, object hit and avoid task, pegboard task	Yes (some measures)	Yes
Naeeri et al. (2021)	Aviation simulator	flight simulator	60	PVT, eye fixation number, eye fixation duration, mean pupil size, visual entropy	Yes (some measures)	Yes
Qin et al. (2021)	Aviation	Simulated flight	90	ECG, Blink rate, Pupil diameter, vigilance, attacking	Yes (some measures)	Yes
Guillemin et al. (2022)	Laboratory	t-load d back	32	blinks, pupil diameter	Yes (some measures)	Yes

Ioannucci (2022a)	Laboratory	Auditory task (n-back, pitch-sequence reproduction or a switch task)	41	Pupil size, EEG, Texture Discrimination Task	Yes (some measures)	Yes
Ioannucci (2022b)	Laboratory	As above	41	Pupil size, Texture Discrimination Task	Yes	Yes
Ioannucci (2022c)	Laboratory	As above	41	Pupil size, Texture Discrimination Task	Yes	N/A
Ioannucci (2022d)	Laboratory	As above	40	Texture Discrimination Task, fMRI	Yes (some measures)	Yes
Lampe & Deml (2022)	Driving Simulator	Driving	40	Eyelid distance, blink frequency, blink duration	Yes (some measures)	Yes
Mehmood et al. (2022)	Workplace	Excavation task	60	Facial features, EEG	Yes (some measures)	Yes
Shi et al. (2022)	Driving Simulator	Driving	90	HR Variability, pupillary light reflex	Yes (some measures)	Yes
Sun et al. (2022)	Workplace	CPT	120	HR Variability, flicker fusion frequency, blink frequency, electrodermal current, pupil diameter, EEG, SART	Yes (some measures)	Yes
Yu et al. (2022)	Laboratory	Go no go task	180	EEG, pupil diameter	Yes (some measures)	Unclear
Zhao et al. (2022)	Driving simulator	Driving	60	Mobile phone call task, Pupil diameter, EEG	Yes (some measures)	N/A

CPT = Continuous Performance Test, HR = Heart Rate, PERCLOS = Percentage of eye closure, PVT = Psychomotor Vigilance Task, fMRI = functional magnetic resonance imaging, EEG = electroencephalogram,

SART = Sustained attention to response task, ECG = electrocardiogram

Objective measures

Many papers recorded task performance data for their cognitive fatigue induction task. These results are not discussed, as these measures would not be suitable for measuring objective cognitive fatigue in naturalistic settings, due to the confound of simultaneously inducing cognitive fatigue.

A total of 191 objective measures of cognitive fatigue were reported across the 57 studies. Sixty-eight of these were only reported in one study, whilst the remaining 123 were reported in at least two. Most could be categorised as either ‘physiological’ (100) or ‘response time’ (50) measures (see Table 3). Eye-related measures (which assessed blinks, eye closures, pupil size and saccades via an eye tracker) were the most frequently used ‘physiological’ measurements (44 studies), followed by EEG (23 studies) and heart rate (21 studies). ‘Response time’ tasks were measured in various ways (including n-back tasks, go no-go tasks, the advanced trail making test, and number-based tasks).

Summary of findings

The full dataset was only available for seven studies, with four more providing data on request. Typically, there was limited discussion of effect sizes (only 15 studies provided effect sizes for all measures; three provided effect sizes for some measures), reliability (only seven mentioned reliability, and rarely in relation to the reliability testing of the measures themselves) or validity (only 11 mentioned validity, and rarely in relation to the validity testing of the measures – three did mention cross validity and two mentioned ecological validity). This and the lack of access to the data limited this review’s second objective (assessment of the reliability and validity of the measures). The lack of reporting (and lack of clarity in reporting) and the heterogeneity of the studies also meant that the planned meta-analysis was not appropriate (Siddaway et al., 2019). As such only a narrative synthesis was conducted.

Riemersma et al. (1978b) was excluded from the narrative synthesis due to unclear reporting of the inferential statistics. Of the remaining 55 studies, 14 evidenced cognitive fatigue after their induction task (i.e., found statistically significant differences pre and post cognitive fatigue induction) across all objective measures, 40 found evidence of cognitive fatigue in at least one objective measure, and 1 found no evidence of objective cognitive fatigue. Importantly, 13 studies did not report some of the measures (35.7% of these related to cognitive tasks, 28.6% related to eye measurements and 7.1% related to other physiological measures).

Eye-related measures and observations of the body (which measured facial features, seat pressure or posture) demonstrated significant differences from baseline most frequently, with neuroimaging and 'other' physiological measures demonstrating this least frequently (see Table 4). This suggests that eye- and body-related measures may be more sensitive measures of objective cognitive fatigue (though when the latter is measured by observations it is important to note that these are susceptible to biases). However, comparisons across measurement types should be cautiously interpreted because eye-related and cognitive measures were far more common (for both there were 65 measures across the studies, compared to two cardiovascular measures, 24 neuroimaging measures and seven ‘other’ physiological measures).

Table 4

Analysis of Objective Measures: Frequency and Proportions of Statistically Significant Differences from Baseline

Measure	**Statistically significant differences**	**No statistically significant differences**	**Some statistically significant differences**	**Inferential test not reported**
Eye-related measures	43 (66.2%)	10 (15.4%)	8 (12.3%)	4 (6.2%)
Observations of the body	3 (75.0%)	0 (0.0%)	1 (25.0%)	0 (0.0%)
Neuroimaging	8 (33.3%)	3 (12.5%)	13 (54.2%)	0 (0.0%)
Cardiovascular (BP, HR)	12 (46.2%)	2 (7.7%)	7 (26.9%)	5 (19.2%)
Other physiological measures	2 (28.6%)	2 (28.6%)	2 (28.6%)	1 (14.3%)
Cognitive tasks	36 (54.5%)	16 (24.2%)	10 (15.2%)	4 (6.1%)

Note, observations of the body included facial features, seat pressure and posture, and ‘other’ physiological measures included EMG, electrodermal responses, oxygen saturation and plethysmography

Two types of neuroimaging were reported: EEG (22 measures) and fMRI (two measures). These results are not considered here due to their current impracticality as 'real-world' measures.

Eye-related measures

Eye-related measures typically focused on three areas; blinks or eye closures (23 measures), eye movements (i.e., saccades; 13 measures) and pupil diameter (including pupil dilation/constriction, pupil size and reflex responses such as those elicited during critical flicker fusion tasks; 29 measures). Typically, the findings demonstrated increased blink frequency and duration, and a higher percentage of eye closure after cognitive fatigue induction or during conditions with higher cognitive load (Gergelyfi et al., 2015; Guo et al., 2018; Guillemin et al., 2022; Herlambang et al., 2019, 2020; Körber et al., 2015; Lampe & Deml, 2022; Qin et al., 2021; Sun et al., 2022; Zargari et al., 2018; Zhao et al., 2017). However, Mckinley et al. (2011) did not find significant changes in the percentage of eye closure after cognitive fatigue, and Lampe & Deml (2022), Hu & Lodewijks (2021a) and Völker et al. (2016) did not find significant differences in blink frequency (though Hu & Lodewijks did find an increase in blink duration). Lampe & Deml (2022) found mixed results with some evidence for greater eyelid distance and some for reduced eyelid distance.

The evidence regarding the effects of cognitive fatigue on pupil size is conflicting, with some finding an increase in pupil diameter, peak dilation and constriction latency (Guillemin et al., 2022; Hopstaken et al., 2015; LeDuc et al., 2005; Zargari et al., 2018), some finding a decrease (Herlambang et al., 2020; Hu & Lodewijks, 2021a; Ioannucci, 2022a; Körber et al., 2015; Ma et al., 2018; Naeeri et al., 2021; Sun et al., 2022), and some finding a decrease followed by an increase (Herlambang et al., 2019). Whilst Qin et al. (2021) found no effect on pupil diameter, and Gergelyfi et al., (2015) found that pupillometry remained constant despite extrinsic reward availability, Hopstaken et al. (2016) found an initial decrease in pupil size that increased after a reward manipulation.

In relation to eye movements, Naeeri et al. (2021) found an increase in entropy (i.e., variability in visual scanning patterns), but McKinley et al. (2011) found reduced entropy and less irregularity in eye movements. Whilst some reported increased saccade frequency and amplitude over time (Herlambang et al., 2019; Hu & Lodewijks, 2021a), others reported a decrease (Herlambang et al., 2020). LeDuc et al. (2005) found decreased constriction amplitude and saccadic velocity, and Zargari et al. (2018) found that saccade duration decreased. Conversely, Hu & Lodewijks (2021a) found some evidence for increased saccade duration and Naeeri et al. (2021) found lower eye fixation frequencies and longer durations during cognitive fatigue. Hopstaken et al. (2016) also found a higher percentage of offscreen/missing gaze position until a reward manipulation when this significantly reduced and gaze towards the task increased. In addition, Herlambang et al. (2019, 2020) reported increased visual distraction frequency and duration over time.

Other eye-related measured focused on light sensitivity. For example, Shi et al. (2022) found changes in pupillary light reflex and acceleration characteristics over time, whereas others found worse performance on critical flicker fusion sensitivity (Feyer et al., 1994; Sun et al., 2022; Williamson et al., 1996), although Williamson et al. showed a linear increase in sensitivity in one of their conditions.

Together, these studies highlight the complex, sometimes contradictory, relationship between cognitive fatigue and eye-related measures. Differing durations of fatigue induction between studies alone do not explain these variations. For return-to-work protocols, eye-related measures offer non-intrusive real-time approaches to measuring cognitive fatigue that enable participants to continue on-task, allowing for naturalistic data unaffected by the issues of self-report. Such measures are also (as evidenced above) sensitive to even subtle changes in cognitive fatigue. However, cognitive fatigue is a multifaceted phenomenon, which eye-related measures alone will not fully capture. Additionally, factors such as lighting, task characteristics and individual differences have the potential to confound eye-related measures, reducing their reliability and validity. Furthermore, eye-related metrics may be sensitive to visual fatigue rather than cognitive fatigue per se (Souchet et al., 2022). A further limitation on applicability for return-to-work protocols is that such measures typically require the user to be static (e.g., at a computer or behind the wheel of a vehicle), although developments in eye tracking glasses may widen their generalisability to other settings.

Observations of the body

Observations of the body were categorised into facial features (two measures), seat pressure (one measure) and posture (one measure). Völker et al. (2016) found that observer-rated facial fatigue increased, and posture sway decreased. Mehmood et al. (2022) found changes within various facial features; decreases in eye aspect ratio and nose-to-chin ratio, and increases in eye distance and eyebrow features, mouth aspect ratio, and mean values of face area and head motion. There were no significant variations in nose-to-jaw ratio. Jagannath & Balasubramanian (2014) found changes in bilateral pressure distribution on thigh and buttock regions; overall mean contact area decreased with cognitive fatigue.

As with eye-related measures, measuring facial features can capture objective 'real-time' cognitive fatigue relatively unobtrusively in naturalistic settings. However, there is typically more subjectivity in the interpretation of facial expressions compared to eye-related measures, and more issues related to cultural and individual differences, which can limit the generalisability, reliability and validity of such measures. Facial features are also less linked to underlying neural or physiological processing so may provide less useful insights compared to eye-related measures. People may also suppress their facial expressions within professional settings, limiting their accuracy. As with eye-related measures, the assessment of facial features or seat pressure would require the user to be relatively static, limiting generalisability.

Cardiovascular measures

There were two types of cardiovascular measures: heart rate (24 studies) and blood pressure (2 studies). Whilst Jagannath & Balasubramanian (2014) reported decreased systolic and diastolic blood pressure during cognitive fatigue, Riese et al., (1999) found no differences. Heart rate was considered in various ways; inter-beat variability, heart rate, and in terms of low, mid or high frequency power (i.e., the distribution of heart rate variability across different frequency bands – high frequency is typically associated with the parasympathetic nervous system).

Whilst Wang et al. (2018) reported reduced heart rate variability during fatigue, most studies found increased inter-beat interval times and variability (Gergelyfi et al., 2015; Riemersma et al., 1978a; Feyer et al., 1994; Williamson et al., 1996; Zhang & Yu, 2010) suggesting decreased alertness. These studies all used fatigue inductions of two hours or more. Using shorter fatigue inductions (of fifty minutes or less), Riese et al. (1999) and Ferreira Junior (2020) found no significant differences in variability, supported by Saito et al. (2018) who suggested no substantial impact on the balance between sympathetic and parasympathetic nervous system activities during a cognitive fatigue-inducing task.

Most studies found a decrease in high frequencies, typically associated with reduced parasympathetic activity and, therefore, increased stress and dominance of the sympathetic nervous system (i.e., Shi et al., 2022; Smith et al., 2019; Tanaka et al., 2009; Tanaka et al., 2011; Zhang & Yu, 2010; Zhang et al., 2018). However, Balasubramanian & Bhardwaj (2018) and Herlambang et al. (2020) found an increase, and Herlambang et al. (2019) found no differences. Herlambang et al. (2019, 2020) observed an increase in midfrequency power over time. No other studies reported any midfrequency power analyses. Most studies found increased low frequency percentages during cognitive fatigue, indicating altered autonomic activity (Qin et al., 2021; Shi et al., 2022; Smith et al., 2019; Tanaka et al., 2011; Zhang & Yu, 2010; Zhang et al., 2018), although Balasubramanian & Bhardwaj (2018) found decreasing trends.

Whilst Balasubramanian & Bhardwaj (2018) found decreasing trends in the low: high frequency ratio after a two-hour fatigue induction, and - with shorter fatigue inductions - Saito et al. (2018), Smith et al. (2019) and Tanaka et al., (2009) found no differences, others observed a linear increase in standard deviation of R–R intervals, and the low: high frequency ratio (Shi et al., 2022; Qin et al., 2021; Zhang & Yu, 2010; Zhang et al., 2018). Additionally, Zhang and Yu (2010) found that total power increased after cognitive fatigue, whereas Zhang et al. (2018) found no significant changes.

Sun et al. (2022) and Van Cutsem et al. (2019) reported increased heart rate after cognitive fatigue; others found a decrease (Jagannath & Balasubramanian, 2014; Wilson et al., 2020; Habay et al., 2021; Li et al., 2004; Pattyn et al., 2008; Zhang & Yu, 2010; Zhang et al., 2018).

Whilst heart rate and blood pressure can provide objective insights (in a non-invasive, real-time manner) into the physiological responses to cognitive fatigue (particularly relating to effects on the autonomic nervous system), they are also multifaceted constructs affected by numerous other factors (e.g., stress, anxiety, physical activity; Tomitani et al., 2020; Torpy et al., 2007; Trapp et al., 2014). This makes it difficult to attribute changes in either measure directly to cognitive fatigue. They are also influenced by individual differences in terms of how blood pressure and heart rate change in relation to different contexts including cognitive fatigue.

'Other' physiological measures

Four types of physiological measures were categorised as 'other'; plethysmography (i.e., changes in blood volume; three studies), electrodermal activity (two studies), electromyography (one study), and oxygen saturation (one study). No differences were reported for oxygen saturation (Jagannath & Balasubramanian, 2014), respiratory parameters or sympathetic tone (although there was some evidence for increased parasympathetic tone; Pattyn et al., 2008). There were mixed results for skin conductance; Sun et al. (2022) found increased electrodermal values after cognitive fatigue, while Gergelyfi et al. (2015) found no significant differences despite the extrinsic rewards available, despite the same duration of fatigue induction.

Whilst plethysmography, electrodermal activity and oxygen saturation (typically measured by pulse oximetry) are relatively non-invasive techniques allowing real-time monitoring, they too are influenced by other factors including individual differences. Electromyography provides a direct measure of muscle activity to describe the impact of cognitive fatigue on motor control, but this is specific to localised muscle fatigue rather than cognitive fatigue *per se*. Additionally, the need for specific electrode placement makes it less practical in real world contexts than some other physiological measures. As electromyography is sensitive to interference from movement it is most useful for static participants.

Cognitive tasks

65 cognitive tasks were used to measure cognitive fatigue, the most common being psychomotor vigilance (13 studies); the remaining measures were used in only one or two studies making it difficult to extract themes. Most studies found that performance worsened after cognitive fatigue induction (demonstrated by increased response times and/or error rates). This was found in a distance estimation task (Liu and Wu 2009; Riemersa et al., 1978a), a colour changing task (Riemersa et al. (1978a), a dot counting task (Riese et al., 1999), a peripheral decision task (Riese et al., 1999), an oddball task (Zhao et al., 2017), a working memory task (Gergelyfi et al., 2015), an object hit and avoid task (Magnuson et al., 2021), a pegboard task (Magnuson et al., 2021), and a visuomotor task (Habay et al., 2021). Van Cutsem et al. (2019) found increased response times for a visuomotor task, but no effect on accuracy; Chen et al. (2021) found reduced accuracy for maths and suspension tasks, but no effect on response times, while Liu and Wu (2009) found slower response times but no difference in performance for arithmetic tasks.

Some studies reported an improvement in performance. Specifically, improvements were noted in digit symbol coding (Jonasson et al., 2018) and symbol searches (Jonasson et al., 2018). Whilst this may indicate practice effects or, perhaps, suggest that these measures are not sensitive enough to detect 'real-time' cognitive fatigue, it is of note that the duration of fatigue induction was thirty minutes, somewhat less than in other studies. There was also evidence of enhanced performance in unstable tracking, albeit with longer response times (Feyer et al., 1994, replicated by Williamson et al., 1996, though with a mid-trip dip in performance), and the Advanced Trail Making Test (Tanaka et al., 2011 though Tanaka et al., 2009 found worse error counts in the Advanced Trail Making Test). Claros-Salinas et al. (2010) found that divided attention response times improved, but accuracy was unaffected. Decision-making was reduced for Fortes et al. (2021), but Feyer et al. (1994) found faster response times with no change in actual decision making. The trade-off between speed and accuracy highlights the importance of considering both aspects in this context.

Some studies showed mixed results whereby performance initially improved at the mid-point, and loss of performance only happened in some conditions, such as within the texture discrimination task from Ioannucci (2022a, 2022b, 2022c). However, given that these findings are all from the same researcher, it may be an artefact of their design.

Some studies reported no significant changes following cognitive fatigue induction. Jonasson et al. (2018) found no effect for attentional blink, divided attention and working memory combined, digit span or Stroop tasks. Similarly, Saito et al. (2018) found no significant changes in a colour response time task. Of note, the duration of fatigue induction in these studies was thirty and forty minutes respectively, somewhat less than in other studies, although other studies with longer fatigue induction durations had similar findings: Feyer et al. (1994) found no significant effect on their vocal response time task, Völker et al. (2016) found no significant differences for number of characters typed or typing errors, and Otto et al. (2018) found no significant differences for their n back task.

For some tasks, the findings were conflicting. For example, whilst performance was reduced in the vigilance task in some studies (Hu & Lodewijks, 2021a and b; Mckinley et al., 2011; Naeeri et al., 2021; Qin et al., 2021; Smith et al., 2019), others found performance was only reduced in some conditions (Li et al., 2004), was improved (Feyer et al., 1994; Williamson et al., 1996) or found no differences (Körber et al., 2015).

Similarly, whilst Guo et al. (2018) found worse performance in the go no-go task, Guo et al., (2015) and Claros-Salinas et al. (2010) only found an effect in response times, with the former only finding this in one condition and the latter finding that accuracy improved; Magnuson et al. (2021) found no effects on accuracy or response times in the go no-go task. Whilst Fortes et al. (2021) and Saito et al. (2018) found no significant changes in the Stroop performance, Fortes et al. found slower response times. Whilst some studies found worse performance on a simple response time task (Claros-Salinas et al., 2010; Ferreira Junior et al., 2020), Gergelyfi et al., (2015) found no effect. Van Cutsem et al. (2019) found increased response times for the flanker task but no effect on accuracy, whereas Habay et al. (2021) found the opposite – no response time effect but accuracy was affected.

Cognitive tasks can mimic real-world tasks and provide a direct assessment of different aspects of objective cognitive fatigue. However, most of these tasks, by their nature, require cognitive resources to complete, which can add to the cognitive fatigue of the participant. Indeed, some tasks that are commonly used to induce cognitive fatigue (e.g., the Stroop, n-back, flanker, go no-go and vigilance/attention-based tasks) were used as measures within some studies. Additionally, it should be noted that using tasks with physical components to assess cognitive fatigue are likely to be affected by physical fatigue as well. Cognitive measures are also difficult to implement within a working day as they involve taking time out to undertake (although sometimes these only take a few minutes). As observed above, practice effects can reduce their validity as measures of objective cognitive fatigue.

It is important to note that the length of the pre- and post-induction cognitive tasks varied across the studies, which may have important implications for future research. Of the 30 studies with pre- and post-induction cognitive tasks, 9 did not detail the length of the tasks. For the remaining studies, the task durations ranged from 2.5 minutes to 30 minutes. Some tasks were consistently timed both pre and post (e.g., ATMT = 30 mins, GoNoGo = 15 mins), whereas other tasks combined multiple shorter tasks. Six tasks were less than five minutes long, seven tasks were exactly five minutes, four tasks lasted between five and ten minutes, and eight were ten minutes or more. Given the small number of studies involved, caution is needed when drawing conclusions about the influence of task length on the results. However, this variation in task length is an important consideration for future research. While shorter tasks are beneficial for naturalistic studies due to their feasibility, they may lack the sensitivity to detect subtle changes in cognitive fatigue. Importantly, most tasks maintained a consistent duration between the pre and post sessions, which helps to reduce variability and ensures that differences in cognitive performance are attributable to the intervention rather than changes in task duration.

Conclusions and future directions

This systematic review highlights the range of measures used to assess objective cognitive fatigue, and the range of settings and populations in which they have been applied. Whilst the overall trend of the findings suggests that most studies successfully induced and measured objective cognitive fatigue, there were some conflicting findings. These could be due to factors including the heterogeneity of the samples and study characteristics, and how cognitive fatigue was operationalised within each study.

The various definitions of cognitive fatigue highlighted the diversity across the studies and the need for a unifying, operationalised definition that encompasses the broad experiences of cognitive fatigue, including subjective and physiological changes and the influence of environmental conditions. The wider synthesis highlighted the need to standardise protocols and report outcomes with clarity and transparency (in line with Open Science practices, e.g., Hagger, 2022) to enable more direct comparisons - the planned meta-analysis was prevented by this lack of clear and detailed reporting (Siddaway et al., 2019).

The current review's main aim was to consider objective measures of cognitive fatigue that could be applied in real-world settings to inform return-to-work protocols. Whilst the physiological measures provide an avenue for the non-invasive and non-intrusive real-time assessment of objective cognitive fatigue, the potential for other factors (e.g., stress or physical exertion) to confound this assessment means that they are unlikely to be appropriate for standalone use within return-to-work protocols, particularly for those with long term physical and/or cognitive impairments. Eye-related measures offer greater resilience and consistency in assessing cognitive fatigue, and several aspects (blink frequency and duration, pupil diameter and saccadic behaviour) could be collected at once to provide a comprehensive picture. However, there are inconsistencies in findings within these measures across studies, so further research is needed to understand these discrepancies. Additionally, such measures are currently most appropriate for relatively sedentary jobs, although technological advances offer promise for their measurement in more active jobs in the future. Awareness of the impact of environmental conditions such as lighting on these measures is required, which could be controlled for to a greater extent in more sedentary jobs.

Cognitive tasks appear to offer the most direct assessment of cognitive fatigue but can potentially contribute to cognitive fatigue itself. Further research is needed into optimising such measures to minimise this confounding effect. The issue of practice effects can also be problematic for some of these measures, so consideration is required as to what kind of cognitive tasks are most appropriate for this purpose, and how the administration of such measures might be optimised to incorporate familiarisation with them beforehand. Based on the literature included in this review, such tasks might include elements relating to vigilance, switching attention and response suppression, might incorporate an interruption element, and have outcomes relating to both response time and accuracy - the findings of reaction time/accuracy studies highlight the need for both aspects to be included in cognitive measures of fatigue. As few studies were conducted over a period equivalent to a working day, further research is needed to assess the reliability of such measures over this timescale. There is some suggestion that some cognitive measures of fatigue may be less likely to identify potential deterioration in performance when the duration of fatigue induction is fifty minutes or less.

Additionally, it is often the case in clinical populations that difficulties do not occur in isolation. Visual/visuo-spatial difficulties are common in people with acquired brain injuries and can be under-reported. A consequence of this would be that visually demanding tasks exacerbate cognitive fatigue due to the additional resources required to compensate. In such instances, it seems less important to distinguish between visual fatigue and cognitive fatigue, as both will be impacting upon real-world performance.

Some measures were excluded for potential unreliability due to being underpowered (making Type 1 statistical errors more likely) or because they were not accompanied by another measure to evidence that objective cognitive fatigue had been induced. Some potentially relevant measures may have been excluded as a result; however, reducing the likelihood of false positives was prioritised given the potential consequences of ill-managed return-to-work protocols. Additionally, 'real-world' data is likely to have a lower signal to noise ratio than laboratory tests given the lack of experimental control in such settings, meaning that the effect sizes of such measures are likely to be reduced.

In the context of return-to-work protocols, concerns can be raised about the validity of responses to self-report and/or behavioural measures, for example if there is the potential of secondary gain. In clinical cognitive assessment, as well as the development of standalone tests of symptom validity, embedded 'effort-testing' has become increasingly routine. Whilst this was not a consideration identified within the current review, it may be that one direction for the future is to attempt to embed a symptom validity related element into the development of objective measures of cognitive fatigue. The majority of the studies did not explicitly assess whether participants were putting forth their best effort on the tasks. This is particularly crucial for cognitive tasks, as cognitive fatigue is measured by performance decline, and if participants are not fully engaged, any observed decline might reflect a lack of motivation or interest rather than true cognitive fatigue, potentially leading to inaccurate conclusions. Some studies attempted to address this issue indirectly. For instance, one study based tasks on practice performance to assess the maximum level of performance for each participant, thereby ensuring engagement. Another study excluded participants with extremely slow response times, suggesting a lack of effort, while others increased motivation by offering rewards or conducting practice sessions to gauge participant ability. Additionally, some studies employed adaptive mechanisms to maintain task difficulty at an appropriate level, such as varying presentation duration or allowing participants to repeat levels in a Stroop task until a required accuracy level was achieved. In a high-load block, one study adjusted a Sudoku puzzle task to generate new puzzles upon completion, fostering long-term engagement. Baseline testing sessions and repeated training were also used in some studies to ensure consistent effort across tasks. In summary, whilst 46 of the 57 studies did not directly measure effort, several employed indirect methods such as practice-based assessments, adaptive task difficulty, and motivation-enhancing strategies to ensure participants were putting forth their best effort. However, future research should explicitly assess participant effort within tasks to mitigate this potential confound in the results.

Given the well-established relationship between sleep and cognitive performance, it is also important to consider how sleep was accounted for within the studies. Unfortunately, this relationship was largely overlooked. Only one study analysed the correlation between sleep and cognitive performance, while another asked participants about their sleep duration prior to the study. A third study examined sleep data separately, concluding that despite participants being night workers, they were effectively managing the demands associated with their job.

Regarding efforts to mitigate the impact of sleep deprivation, several studies took measures to control for sleep-related variables. For example, some studies provided participants with time to sleep between driving runs or counterbalanced the order in which participants slept. Others obtained detailed information about participants' sleep patterns and sleepiness or required participants to be well-rested before the experiment. Several studies also conducted tests at consistent times of the day to reduce circadian rhythm effects, and some instructed participants to maintain regular sleep schedules or refrain from consuming caffeine or alcohol before testing. However, it is notable that a significant number of studies (21) did not mitigate for sleep deprivation in any way, potentially impacting the validity of their findings.

A secondary consideration for this review was subjective cognitive fatigue. Of the studies included, 48 reported subjective measures, with 39 of these reporting statistically significant differences pre- and post- fatigue induction. In 7 studies it was unclear whether there were differences, and in 2 there were differences in some but not all subjective measures. Thus, a further area for research is on assessing the reliability of subjective measures of cognitive fatigue.

Over 200 articles were inaccessible despite requests directly to authors, which may have contained eligible measures. The lack of accessibility of the data also prevented a meta-analysis from being completed to scrutinise the findings further. This highlights the need for more Open Science practices within the field.

In conclusion, this systematic review highlights the need for a unifying definition of cognitive fatigue, to facilitate measurement that is comprehensive enough to capture the construct more fully. The data synthesis revealed the lack of consistency within findings in the literature, which is likely to be related, at least in part, to a lack of transparency in reporting, and heterogeneity in sample and study characteristics. Eye-related and cognitive measures appear to hold the most promise for monitoring objective cognitive fatigue in real-time in 'real world' settings. However, further research is required relating to cross-population differences and the impacts of the contextual situation. This research is important to inform return-to-work protocols for those returning from illness or injury and should be a priority as part of a wider protocol to improve mental health (especially within the workplace). Objective measures that can inform structured fatigue management planning have the potential to support individuals with long-term conditions (or persisting consequences of acute conditions), their families, colleagues, employers and health professionals within multidisciplinary teams in implementing achievable and sustainable graded returns to activity.

Acknowledgements

This work was supported by an internal grant from the University of Gloucestershire as part of an MSc by Research in Psychology.

We thank Professor Graham Edgar for discussion on key ideas presented in this manuscript.

Declaration of interest statement

The authors report there are no competing interests to declare.

Data Availability Statement

The authors confirm that the data supporting the findings of this study are available within the supplementary materials.

References

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

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

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

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*, 15. https:// doi.org/10.3389/fnhum.2021.703422

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

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

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

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

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. https://doi.org/10.3758/ BF03193146

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

Fortes, L. S., Gantois, P., de Lima-Júnior, D., Barbosa, B. T., Ferreira, M. E. C., Nakamura, F. Y., Albuquerque, M. R., & Fonseca, F. S. (2023). Playing videogames or using social media applications on smartphones causes mental fatigue and impairs decision-making performance in amateur boxers. *Applied Neuropsychology*. 30(2), 227–238. https://doi.org/10.1080/23279095.2021.1927036

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, 9. https://doi.org/10.3389/ fnbeh.2015.00176

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, W., Ren, J., Wang, B., & Zhu, Q. (2015). Effects of relaxing music on mental fatigue induced by a continuous performance task: Behavioral and ERPs evidence. *PLOS ONE*, 10(8), e0136446.

Guo, Z., Chen, R., Liu, X., Zhao, G., Zheng, Y., Gong, M., Zhang, J., & Rao, H. (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., & DiRusso, F. (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., Van Cutsem, J., & 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. https://doi.org/10.3390/ ijerph182412906

Hagger, M. S. (2022). Developing an open science 'mindset'. *Health Psychology and Behavioral Medicine*, 10(1), 1–21. https://doi.org/10.1080/21642850.2021.2012474

Herlambang, M. B., Cnossen, F., Taatgen, N. A., & Steinborn, M. B. (2021). The effects of intrinsic motivation on mental fatigue. *PLOS ONE*, 16(1), e0243754. 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 & Ergonomics Society*, 61(7), 1171–1185. https://doi.org/10.1177/ 0018720819828569

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V. (2019). Cochrane handbook for systematic reviews of interventions (2nd ed.). John Wiley & Sons.

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

Hopstaken, J. (2016). Conquering fatigue: The battle for engagement. Erasmus university Rotterdam. Investigations | diagnosis | tiredness/fatigue in adults | CKS | NICE. (N.D.). *JOURNAL OF CLINICAL AND EXPERIMENTAL NEUROPSYCHOLOGY* 15 Retrieved December 14, 2021, from https://cks.nice.org.uk/ topics/tiredness-fatigue-in-adults/diagnosis/investigations/

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

Hu, X., & Lodewijks, G. (2021). Exploration of the effects of task-related fatigue on eye-motion features and its value in improving driver fatigue-related technology. *Transportation Research. Part F, Traffic Psychology and Behaviour,* 80, 150–171.

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

Joanna Briggs Institute. (2014). Joanna Briggs institute reviewers' manual: 2014 edition. The Joanna Briggs Institute.

Johansson, 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). https://doi.org/10.1002/brb3.1056

Johansson, B., & Rönnbäck, L. (2012). Mental fatigue and cognitive impairment after an almost neurological recovered stroke. ISRN Psychiatry, 2012, 1–7. https://doi.org/10.5402/2012/686425

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

Lampe, D., & Deml, B. (2022). Reducing passive driver fatigue through a suitable secondary motor task by means of an interactive seating system. *Applied Ergonomics*, 103, 103773.

LeDuc, P., Greig, J., & Dumond, S. (2005). Self-report and ocular measures of fatigue in U.S. Army apache aviators following flight | semantic scholar. https://www.semanticscholar.org/paper/Self-Report-and-Ocular-Measures-ofFatigue-in-U.S.-LeDuc-Greig/8a8529433d26741 f8ef333d1a205619a75a359a6

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

Liu, Y. C., & Wu, T. J. (2009). Fatigued driver's driving behavior and cognitive task performance: Effects of road environments and road environment changes. *Safety Science*, 47(8), 1083–1089.

Losoi, H., Silverberg, N. D., Wäljas, M., Turunen, S., RostiOtajärvi, 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

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

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: The Journal of the Human Factors & Ergonomics Society*, 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

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

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, 8(5), 336–341. https://doi.org/10.1016/ j.ijsu.2010.02.007

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), 283. https://doi. org/10.3390/aerospace8100283

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

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 16

Otto, T., Zijlstra, F. R. H., Goebel, R., & Verguts, T. (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), e0198204. 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

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

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

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

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*, 9, 0. https://doi.org/10.3389/fneur.2018.01165

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. 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

Riva, J. J., Malik, K. M. P., Burnie, S. J., Endicott, A. R., & Busse, J. W. (2012). What is your research question? An introduction to the PICOT format for clinicians. *The Journal of the Canadian Chiropractic Association*, 56(3), 167–171. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3430448/

Rönnbäck, L., & Johansson, B. (2022). Long-lasting pathological mental fatigue after brain injury–A dysfunction in glutamate neurotransmission? *Frontiers in Behavioral Neuroscience*, 15. https://doi.org/10.3389/fnbeh.2021. 791984

Saito, N., Yamano, E., Ishii, A., Tanaka, M., Nakamura, J., Watanabe, Y., & Roman, G. (2018). Involvement of the olfactory system in the induction of anti-fatigue effects by odorants. *PLOS ONE*, 13(3), e0195263. https://doi.org/10. 1371/journal.pone.0195263

Shi, L., Zheng, L., Jin, D., Lin, Z., Zhang, Q., & Zhang, M. (2022). Assessment of combination of automated pupillometry and heart rate variability to detect driving fatigue. *Frontiers in Public Health*, 10, 828428. https://doi.org/10. 3389/fpubh.2022.828428

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, 70(1), 747–770. https://doi.org/10.1146/annurev-psych-010418- 102803

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., Tajima, S., Sasabe, T., & Watanabe, Y. (2009). Central nervous system fatigue alters autonomic nerve activity. *Life Sciences*, 84(7–8), 235–239. https://doi.org/10.1016/j.lfs.2008.12.004

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(1), 46. https://doi.org/10.1186/ 1744-9081-7-46

Tomitani, N., Kanegae, H., Suzuki, Y., Kuwabara, M., & Kario, K. (2021). Stress-induced blood pressure elevation self-measured by a wearable watch-type device. *American Journal of Hypertension*, 34(4), 377–382. https://doi.org/10. 1093/ajh/hpaa139

Torpy, J. M., Burke, A. E., & Glass, R. M. (2007). Acute emotional stress and the heart. *JAMA*, 298(3), 360. https://doi.org/10.1001/jama.286.3.374

Trapp, M., Trapp, E.-M., Egger, J. W., Domej, W., Schillaci, G., Avian, A., Rohrer, P. M., Hörlesberger, N., Magometschnigg, D., Cervar-Zivkovic, M., Komericki, P., Velik, R., Baulmann, J., & Torrens, C. (2014). Impact of mental and physical stress on blood pressure and pulse pressure under normobaric versus hypoxic conditions. *PLOS ONE*, 9(5), e89005. https://doi.org/10.1371/journal. pone.0089005

Van Cutsem, J., De Pauw, K., Vandervaeren, C., Marcora, S., Meeusen, R., & Roelands, B. (2019). Mental fatigue impairs visuomotor response time in badminton players, *JOURNAL OF CLINICAL AND EXPERIMENTAL NEUROPSYCHOLOGY*

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

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. https://doi.org/10.3390/e20030196

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 & Ergonomics Society*, 62(4), 553–564. https://doi. org/10.1177/0018720819849783

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*, 12(1), 846. https://doi.org/10.1038/ s41598-022-04972-y

Zargari Marandi, R., Madeleine, P., Omland, Ø., Vuillerme, N., & Samani, A. (2018). Eye movement characteristics reflected fatigue development in both young and elderly individuals. Scientific Reports (Nature Publisher Group), 8(1). 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. https://doi.org/10.2478/v10013- 010-0007-7

Zhang, C., Zheng, C. X., & Yu, X. L. (2009). 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.

Zhao, C., Zhao, M., Yang, Y., Gao, J., Rao, N., & Lin, P. (2016). The reorganization of human brain networks modulated by driving mental fatigue. *IEEE Journal of Biomedical and Health Informatics*, 21(3), 743–755. https://doi.org/10.1109/JBHI.2016.2544061

Zhao, H., Ma, J., Zhang, Y., & Chang, R. (2022). Mental workload accumulation effect of mobile phone distraction in L2 autopilot mode. *Scientific Reports*, 12(1). https://doi.org/10.1038/s41598-022-17419-1