2021

Physiological Measurements for Real-time Fatigue Monitoring in Train Drivers: Review of the State of the Art and Reframing the Problem

Bojana Bjegojevic

Maria Chiara Leva

Nora Balfe

See next page for additional authors

Authors

Bojana Bjegojevic, Maria Chiara Leva, Nora Balfe, Sam D. Cromie, and Luca Longo

Physiological Measurements for Real-time Fatigue Monitoring in Train Drivers: Review of the State of the Art and Reframing the Problem

Bojana Bjegojević

School of Environmental Health, Technological University Dublin, Ireland. E-mail: 20125651@mytudublin.ie

Maria Chiara Leva

School of Environmental Health, Technological University Dublin, Ireland. E-mail: mariachiara.leva@tudublin.ie

Nora Balfe

Health and Safety Directorate, Irish Rail, Ireland. E-mail: Nora.Balfe@irishrail.ie

Sam Cromie

Centre for Innovative Human Systems, School of Psychology, Trinity College Dublin, Ireland. E-mail: sdcromie@tcd.ie

Luca Longo

School of Computer Science, Technological University Dublin, Ireland. E-mail: Luca.Longo@tudublin.ie

The impact of fatigue on train drivers is one of the most important safety-critical issues in rail. It affects drivers' performance, significantly contributing to railway incidents and accidents. To address the issue of real-time fatigue detection in drivers, most reliable and applicable psychophysiological indicators of fatigue need to be identified. Hence, this paper aims to examine and present the current state of the art in physiological measures for real-time fatigue monitoring that could be applied in the train driving context. Three groups of such measures are identified: EEG, eye-tracking and heart-rate measures. This is the first paper to provide the analysis and review of these measures together on a granular level, focusing on specific variables. Their potential application to monitoring train driver fatigue is discussed in respective sections. A summary of all variables, key findings and issues across these measures is provided. An alternative reconceptualization of the problem is proposed, shifting the focus from the concept of fatigue to that of attention. Several arguments are put forward in support of attention as a better-defined construct, more predictive of performance decrements than fatigue, with serious ramifications on human safety. Proposed reframing of the problem coupled with the detailed presentation of findings for specific relevant variables can serve as a guideline for future empirical research, which is needed in this field.

Keywords: Train drivers, Rail, Physiology, Fatigue, Attention, EEG, Eye-tracking, Heart rate.

1. Introduction

The safety implications of the European rail network are becoming pivotal due to the increase in passenger numbers witnessed in recent years (Sangiorgio et al., 2020). The current year of 2021 has been declared as 'the European year of rail' by the European Commission initiative promoting trains in favor of other means of transport in order to reduce greenhouse gas emissions. In addition, the SARS-CoV-2 pandemic, currently reshaping the global travel and transportation landscape as we know, could lead to a further increase in the use of railway in both national and international post-pandemic-era travel. To accommodate the expected upcoming augmentation in rail travel demands, the railway industry needs to ensure it maintains its status as one of the safest means of transportation (ERA, 2020, 27).

Train drivers are one of the most important safety-critical roles in railway. A substantial amount of research has been conducted on the human factors involved in it, especially fatigue (e.g. Branton, 1979; Hamilton & Clarke, 2005; Naweed, 2014; McLeod et al., 2005). Fatigue is regarded as a major risk in all transport industries due to its strong link with performance impairment (Horrey et al., 2011). The UK Rail Accident Investigation Board identified 74 accidents and incidents between 2001 and 2009 in which fatigue was a contributory factor (Young & Steel, 2016). Furthermore, train drivers have very irregular working hours and they cannot take a break whenever they need it (Kecklund et al. 1999), which can increase their fatigue levels. A recent study (Fan & Smith, 2020) demonstrated the association between work-related fatigue and impaired performance in train drivers and other railway staff. Restricted train maneuverability, significant amount of action planning, and long monotonous periods with minimum required activity have also been recognized as major risks for instigating fatigue and reducing alertness or vigilance in train drivers (Filtness & Naweed, 2017).

To manage fatigue, it is of essence to identify it promptly i.e. in real time, and objectively. A growing number of studies are using various psychophysiological

Proceedings of the 31st European Safety and Reliability Conference
Edited by Bruno Castanier, Marko Cepin, David Bigaud, and Christophe Berenguer

ISBN: 978-981-18-2016-8; doi:10.3850/978-981-18-2016-8_437-cd

measures for detecting fatigue in drivers, as they have three main advantages: 1) They provide continuous rather than discrete measurements, which is a prerequisite for real-time fatigue monitoring; 2) They do not interrupt the process of driving unlike the inclusion of secondary tasks, nor do they exert additional time and energy resources, like the use of Psychomotor Vigilance Task; 3) They are not subject to manipulation as are self-assessment methods, allowing for more objective comparison. Hence, this review aims to examine the current state of the art in physiological measures for real-time fatigue monitoring that could be applied in train driving context.

2. Defining the scope and content of literature review

Methods for real-time fatigue monitoring identified through the first stage of literature search encompass electroencephalography (EEG), eye-tracking (ET), heart-rate- (HR), galvanic skin response- (GSR) also known as electrodermal activity- (EDA) monitoring, electromyography (EMG), and video-recordings[a]. Having in mind the applicability to train driving, several of the mentioned methods have been excluded from further research. Due to the nature of hand- and body- movements involved in train driving, neither GSR nor video-recordings could be applied. GSR is most accurately measured from the fingertips, but placing the sensors on them is obtrusive to the driving task and the signals would be affected by any pressure involved in button-pressing (Crowley & Balfe, 2018). Wrist-based devices for EDA/GSR monitoring are at present not reliable enough for field studies (Mehler et al., 2018). Train drivers have to frequently stand up and look outside the windows which also does not allow for the use of remote-based devices (detached from the driver) such as stationary eye-trackers and video-cameras. Video recordings also impose privacy issues and have the potential to cause mental discomfort. EMG is most useful in examining muscular or physical fatigue (e.g. Jagannath & Balasubramanian, 2014). Therefore, it is less suited for assessing train drivers' fatigue, as train driving is predominantly a sedentary type of work. Thus, three remaining groups of measures for real-time fatigue assessment in train drivers are reviewed in this paper. These are electroencephalography (EEG), eye-tracking (ET), and heart-rate monitoring (HR), and are presented in respective sections below. Relevant findings and physiological variables are identified and discussed. For purpose of clarity, an overview of the steps involved in the review of the state of the art is presented below (Figure 1).

[a] Although video recordings are not physiological measurements, they do provide real-time and objective data.

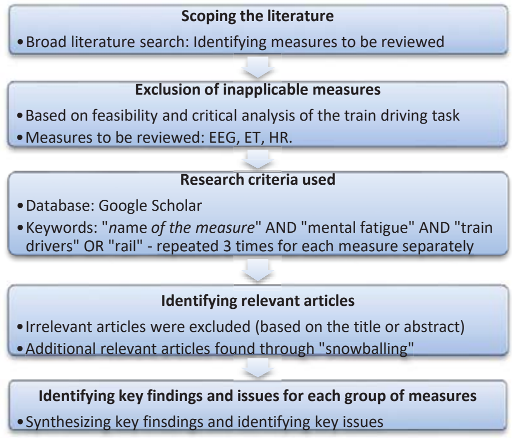

Fig. 1: Steps followed in the current review of the state of the art.

2.1. *EEG measures*

Electroencephalography (EEG) is a non-invasive method for measuring electrical brain activity originating mostly in the cortex. Electrical signals are registered with multiple electrodes, usually embedded in a cap or a net, placed on the scalp. EEG signals are most commonly analysed in the frequency or time-frequency domain (spectral-power analysis) and in the time domain (Event-Related Potential or ERP analysis) (Gramann & Plank, 2019).

2.1.1. *Spectral-power measures (frequency bands)*

Most of the spectral-power-based studies focused on the effect of fatigue on the changes in the alpha, theta, beta, delta, or derived indices from some combination of those bands (Lees et al., 2018; Jap et al., 2011; Jap et al., 2007; Lal & Craig, 2002; Lal & Craig, 2001; Lal & Craig 2000). Summarizing all the findings relating EEG activity to fatigue is challenging, as the studies vary greatly, even within the driving application context. The number of studies on this topic has significantly grown in recent years and the majority focused on employing machine learning algorithms to detect the occurrence of fatigue in drivers, with various degrees of success and variability in study design, measurement and analyses (see Sahayadhas et al., 2012). They often do not report the direction of change in the parameters used, and focus on obtaining the highest possible accuracy of classification of fatigued states (e.g. Guarda et al., 2018) which are diversely defined. This approach can provide applicable results only when the basic understanding of the physiology of the fatigued brain is established. Therefore, we will present here only the papers that clearly indicated the direction of change in specific frequency band powers that occurred as a result of fatigue.

According to Lal and Craig who have extensively researched the relationship between driver fatigue and physiology, "the literature abounds with EEG changes associated with fatigue but the results are quite variable"

(Lal & Craig, 2001, 182). They state that increase in alpha and theta, and reduction of beta waves are considered indicators of sleepiness and fatigue. However, in their later study (Lal & Craig, 2002) they found an increase in theta and delta activity, but also an increase in alpha and beta, albeit smaller. With an increase of fatigue, a small increase in beta along with a more prominent increase in theta and lower-alpha band was also found in a controlled experimental study combining EEG and ERP analyses (Boksem et al., 2005). A potential explanation of such a pattern of results regarding beta is offered by Craig et al. (2012). Namely, with increasing fatigue, the activity of the brain slows down, so the fatigued individual increases their efforts to sustain vigilance, which in turn leads to increase in beta activity. A study on train drivers conducting a monotonous driving task in a simulator showed simultaneous decreases in beta and increases in theta, as the train drivers were becoming fatigued (Jap et al., 2011), which is in line with the expected pattern, as stated earlier. They also found that delta activity had started to increase and then fluctuated. In the same study, the authors compared various combinations of EEG indices attempting to determine which of them differentiates best between the alert and fatigued state. They concluded it to be the sum of theta and alpha divided by beta. Another study investigated the relationship between EEG activity of train drivers performing a 10-minute monotonous simulated driving task and self-report fatigue and sleepiness scores (Lees et al., 2018). Delta, theta, and alpha were found to be negatively correlated with those scores. However, fatigue self-assessment was conducted prior to simulated driving and there was no experimental manipulation of fatigue. Therefore, these results are not informative of the changes in the EEG influenced directly by fatigue.

In summary, not many studies have empirically demonstrated the relationship between specific EEG bands and fatigue. Researchers frequently *a priori* assume that different EEG frequencies or their combinations are related to fatigue (e.g. Jap & Lal, 2011), and that is probably part of the reason why there are many conflicting or unclear findings in this field.

2.1.2. *ERPs*

The ERP-based studies require a much more controlled experiment design and have low ecological validity. At this stage, the ERP method cannot be used in real-time fatigue monitoring. However, some of the most relevant conclusions of the ERP-based studies are mentioned here briefly as they provide deeper insight into the cognitive processes affected by fatigue. For instance, Boksem et al. (2015) have demonstrated a differential effect of fatigue over P1, N1, and N2b ERP components. This result suggests that the top-down (goal-directed) aspect of attention is adversely impacted by fatigue, whereas bottom-up (stimulus-driven) attentional processes are not. This agrees with the conclusion of an earlier study (Faber et al., 2012) that fatigued individuals are more susceptible to distractions and less able to suppress irrelevant signals, which can lead to more errors. Such findings are important in understanding which kinds of safety-critical tasks suffer most from the occurrence of fatigue.

EEG is known to possess excellent temporal resolution compared to other neuroimaging techniques, but less precise spatial resolution. This overview has therefore focused solely on the direction of changes in frequency bands, disregarding the electrode sites. The lack of a standard experimental and reporting protocol also makes the attempt of summarising localisation results extremely challenging. Furthermore, different authors define the ranges of EEG bands differently, and the same issue occurs in defining fatigue, which can cause additional problems in the interpretation of the findings.

Although EEG is often referred to as the ‘gold standard’ for assessing mental fatigue (e.g. Liu et al., 2010), there is an obvious need for thorough and large-scale systematization of various EEG findings relevant to fatigue and (train) driving. Future studies in this field should be more theory-based i.e. conclusions should not be driven solely by obtained data; they should be more methodologically rigorous; and the study design and analyses procedures should be reported thoroughly.

2.2. *Eye-tracking measures*

There are number of measures that can be obtained from monitoring the eyes. These measures are here divided into three groups: pupil-, blink-, and eye-movement-based measures. They will be discussed below in that order.

2.2.1. *Pupil-based variables*

Pupillometry is the term used for measuring changes in the pupil diameter. These changes serve to regulate the amount of light passing through the eyes: as the amount of light increases, the pupil size decreases. However, luminance is not the only factor that affects the size of pupil diameter; it is also influenced by attention and cognition (Peysakhovich et al., 2015), as well as emotional arousal (Bradley et al., 2008).

Changes in pupil size are involuntary reactions that cannot be controlled (Leang et al., 2012). This qualifies them as an objective indicator of various physiological states, including fatigue. Pupil size was shown to decrease with time-on-task (Hopstaken et al., 2016). Moreover, when a person is experiencing fatigue, pupil size tends to decrease and have greater variability (Brisson et al, 2013 & Gilzenrat et al, 2010, as cited in Sirois & Brisson, 2014). The bigger variability of pupil size during fatigue is a consequence of the competition between the opposing effects of the sympathetic and parasympathetic nervous systems (Straeter, 2020).

The challenge in the use of pupillometrics as an objective indicator of physiological state lies in the fact that pupil dilation and constriction can occur as a result the factors mentioned earlier (light conditions, mental effort, emotional arousal). The changes in pupil size that occur due to cognitive factors, such as mental fatigue, are small compared to those evoked by light conditions (Laeng et al., 2012). To circumvent some of those obstacles, Straeter (2020) proposed using the gliding average method for calculating the changes in pupil size over time, which mathematically controls for the pupil diameter variation caused by the luminance changes. With the use of such corrections, pupillometry seems as a promising method for monitoring mental fatigue, especially since it is not controlled voluntarily.

2.2.2. *Blink-based variables*

The second group of measures comprises various blink parameters. Many studies assume the existence of a direct relationship between blinks and fatigue or drowsiness (e.g. Hsieh & Tai, 2013). A systematic review (Martins & Carvalho, 2015) summarized papers from 2010 onwards that used eye blinks as indicators of 'fatigue or mental load'. However, it is not clear from the review how successful these studies were in detecting fatigue or mental load based on blinking patterns. The only measure reported was blink frequency. In one of the reviewed studies (Caffier et al., 2003), blink frequency or rate was shown to slightly decrease when participants were tired. On the contrary, another, more recent study (Maffei & Angrilli, 2018) found an increase in blink rate over time-on-task, which the authors interpret as arousal decrement and an increase in mental fatigue. More importantly, they found that blink rate is affected both by task difficulty and time-on-task: blink rate increased over the course of easier tasks, but not over the course of a difficult task. The authors claim that blinking is supressed in difficult tasks as they require more attention, which confirms their validity as an index of (sustained) attention.

Apart from blink frequency, blink duration can also be indicative of tiredness and fatigue. In one study, mean blink duration significantly increased from morning/alert compared to evening/drowsy condition, correlating positively with the subjective estimation of drowsiness (Caffier et al., 2003). Another study (McIntire et al., 2014) showed that as the performance declines, blink frequency and duration increase. Morris and Miller (1996) found the blink amplitude to be the best predictor of performance decrements due to fatigue, followed by blink rate, long closure rate, and blink duration; and proposed combining them to explain the maximal amount of variance in predicting the performance outcome.

The percentage of eye closures, widely known as PERCLOS is a derived measure often used in assessment of driver fatigue. It was shown to increase as the levels of fatigue rise (Wang et al., 2010). Although PERCLOS was claimed to be a superior indicator of fatigue, Trutschel et al. (2011) found other measures (EEG and eye-tracking signals) more informative.

Since blinking can be partially controlled by will, and is affected by time on task (Maffei & Angrilli, 2018; Stern et al., 1994), time of the day (Barbato et al, 2000, as cited in Caffier et al., 2003), and temperature and humidity (Martins & Carvalho, 2015), blink-based measures do not seem reliable enough for fatigue assessment.

2.2.3. *Eye-movement-based variables*

The third group of measures is based on the monitoring the eye movements. For purpose of clarity, eye movements will be categorised here into fixations and saccades.

Fixations represent the static gaze on a specific area with the duration of 200-300 milliseconds that are separated by saccades (Shrestha & Owens, 2008). There are not many papers that have investigated the relationship between eye-fixation variables and fatigue. One of the few exceptions is a study by Bodala et al. (2017). They devised a fixation score measure which can be used as an indicator of vigilance decrement with the aid of computer vision assisted eye-tracking. It basically shows how much the gaze deviates from the target area of interest.

The ballistic movement of the eyes from one point of fixation to another are known as saccades (Di Stasi et al., 2013). There are many different parameters that can be retrieved from analyzing saccades, the most common ones being their amplitude and velocity. Both were shown to decrease as the vigilance drops (Bodala et al, 2016). Saccadic peak velocity is considered a good index of arousal (Di Stasi, 2013). As opposed to saccadic amplitude and fixation duration, saccadic velocity is not subject to voluntary control (Leigh & Zee, 1999, as cited in Di Stasi et al., 2013), which qualifies them as a more objective indicator of fatigue.

Various devices can be employed for recording ocular measures. Basic variables such as blink occurrence and duration or vertical and horizontal eye movements can be obtained with the use of electrooculogram (EOG) or even cameras. However, for more precise measurements and more promising variables for fatigue monitoring (such as fixations and saccades), eye-trackers are required.

2.3. *Heart rate measures*

Similar to EEG, heart rate signals can be analysed both in the time and the frequency domains, but also using other methods (for detailed overview of these methods and variables see Malik et al., 1996; Shaffer & Ginsberg, 2017). Listing all these variables would exceed the scope of this paper, so we will focus here only on basic heart rate and heart rate variability.

2.3.1. *Heart-rate (HR)*

With regards to the basic variable of heart rate and its relation to fatigue, not many relevant studies have been found. In one study assessing the onset of driver fatigue in a simulator, HR decreased significantly during the course of monotonous driving (Jagannath & Balasubramanian, 2014).

2.3.2. *Heart-rate variability (HRV)*

Heart rate variability is a more complex variable of cardiac activity that might be used as an indicator of fatigue. It is considered a measure of neurocardiac function as it mirrors the interaction between the brain and the heart (Shaffer et al., 2014). Basically, it represents "the changes in the time intervals between consecutive heartbeats called interbeat intervals (IBIs)" (Shaffer & Ginsberg, 2017, 1). It is defined as "an RR interval (RRI) fluctuation on an electrocardiogram trace" (Fujiwara et al., 2019, 1769), measured in milliseconds. The Rs represent the highest peaks on the ECG trace in each cardiac cycle. It seems that RR intervals are used interchangeably with the NN intervals by many authors (e.g. Shaffer & Ginsberg, 2017), which are the intervals between normal R-peaks only (not affected by signal noise or underlying heart conditions). This is important as the Standard Deviation of NNs (SDNN) is a recommended time-domain variable to measure, according to European Heart Journal guidelines (Malik et al., 1996). Two other spectral-domain variables are also recommended: low and high frequency range of heart signals (LF and HF). Their ratio (LF/HF) is understood to be an indicator of the ratio between the sympathetic and parasympathetic nervous system activity (Shaffer & Ginsberg, 2017). This could be a useful parameter as the sympathicus is responsible for the "fight or flight" response activation, indicating the presence of a stressful situation, while the parasympathicus is a dampening system activated in the "rest and digest" or stress-free conditions (LeBouef et al., 2020). Not many literature resources found were clear on which of these HRV variables were used. A recent study investigating the relationship between EEG and ECG signals is an exception (Lees et al., 2021). In that study, an association was established between LF HRV and increased delta and decreased beta during simulated monotonous train driving.

Heart rate measures are most precisely obtained with ECG electrodes, which is impractical in the context of train driving. Wearable devices (smart watches, fitness bands) are currently not precise enough for HRV analysis (Fujiwara et al., 2019). However, chest straps for ECG measurements might prove reliable enough for this purpose.

3. Discussion of the literature findings

The main findings of our review are summarized in Table 1. All three groups of measures are evidently complex and a great number of variables can be defined within each of them. This 'microscopic' view serves to indicate not only the potentially relevant paths for future research, but also the heterogeneity and complexity of the phenomena.

Although many of the reviewed variables seem promising, none of them have been demonstrated to be adequate indicators of fatigue that can be employed in real time monitoring of train drivers at present. This conclusion is in line with a statement from a recent article by Yuzhong et al. (2020, 560): "there are still many problems in the fatigue detection technology based on the physiological characteristics of the workers, and it is not possible to achieve a high accuracy rate in a short time".

We argue that one of the main issues behind the inconclusiveness of the presented findings is the heterogeneity of the concept of fatigue. It is often used synonymously with the terms 'drowsiness' or 'sleepiness' (e.g. Lees et al., 2018), indicating a lack of a precise operational definition. Many of the reviewed papers do not provide a definition of fatigue, and most that do, commonly define it through the concepts of attention and performance. For that reason, a suggestion to reframe and redefine the problem of real-time fatigue monitoring in train drivers is put forward. We instead propose the assessment of attention levels through the use of physiological parameters.

Table 1. Summary of key findings and issues

Physiological measure	Key findings and issues
EEG measures: frequency bands	**<u>Key findings:</u>** • Increase in theta *(Lal & Craig, 2002; Boksem et al., 2005; Jap et al., 2011)* • Increase in delta *(Lal & Craig, 2002);* followed by fluctuation *(Jap et al., 2011)* • Increase in alpha *(Lal & Craig, 2002)* • Small increase in beta (*Lal & Craig, 2002; Boksem et al., 2005*); decrease in beta *(Jap et al., 2011)* • Combinations of these frequencies, e.g. (theta + alpha) / beta, may be more predictive of fatigue *(Jap et al., 2011)* **<u>Key issues:</u>** • Lack of sufficient empirical evidence for the relationship between specific EEG bands and specific mental states, such as fatigued state • Research relying *a priori* on the assumption that different EEG frequencies or their combinations are related to fatigue (e.g. *Jap & Lal, 2011*)
EEG measures: ERP	**<u>Key findings:</u>** • Goal-directed attention is adversely impacted by fatigue, unlike the stimulus-driven attention *(Boksem et al., 2015)* • Fatigued individuals are more susceptible to distractions and less able to suppress irrelevant signals *(Faber et al., 2012)* **<u>Key issues:</u>** • Low ecological validity; not applicable for continuous monitoring
Pupil-based measures	**<u>Key findings:</u>** • Decrease and greater variability of pupil size *(Sirois & Brisson, 2014)* **<u>Key issues:</u>** • Luminance affects the size of pupil; correction method needs to be applied *(Straeter, 2020)*
Blink-based measures	**<u>Key findings:</u>** • Slight decrease in blink rate of tired subjects *(Caffier et al., 2003)*; blink rate increase as performance declines *(McIntire et al., 2014)* and over the course of easier but not difficult tasks due to attention allocation *(Maffei & Angrilli, 2018)* • PERCLOS increases with fatigue *(Wang et al., 2010)*; other measures more informative *(Trutschel et al., 2011)* • Blink amplitude, rate, long closure rate and blink duration combined explain the maximum amount of performance variance *(Morris & Miller, 1996)* **<u>Key issues:</u>** • Blinking can be controlled at will • It is heavily affected by time on task *(Maffei & Angrilli, 2018; Stern et al., 1994)*, time of the day *(Caffier et al., 2003)*, temperature and humidity *(Martins & Carvalho, 2015)*, and task difficulty *(Maffei & Angrilli, 2018)* • More suitable as an index of attention *(Maffei & Angrilli, 2018)*

Eye-movement-based measures	**Key findings:** • Fixation score can be used as an indicator of vigilance decrement *(Bodala et al., 2017)* • Saccadic amplitude and velocity decrease as vigilance drops *(Bodala et al, 2016)*; Saccadic peak velocity is a good index of arousal *(Di Stasi, 2013)* **Key issues:** • Not many papers on the relationship between eye-fixation variables and fatigue • Saccadic amplitude and fixation duration are subject to voluntary control *(Di Stasi, 2013)*
Heart-based measures: HR	**Key findings:** • Decrease during the course of monotonous driving *(Jagannath & Balasubramanian, 2014)* **Key issues:** • Not enough empirical studies have been found
Heart-based measures: HRV	**Key findings:** • Potentially useful variables to explore further in future studies: SDNN; LF and HF, and their ratio **Key issues:** • Not always clear in the literature which HRV method of calculation and variables were used

4. Proposed reframing of the problem: From attempting to measure fatigue to measuring attention

Four main arguments are presented here to support our proposed shift to investigation of the concept of attention. The first is related to the lack of clear operational definition of fatigue, which seems to be a quite diversely defined concept. This is further supported by Williamson et al. statement (2011, 498): “Fatigue is a hypothetical construct (. . .) inferred because it produces measurable phenomena even though it may not be directly observable or objectively measurable”. Narrowing down the literature review to mental fatigue still suffers from lack of operationalisation. Papers on mental fatigue (e.g. Shen et al., 2008; Lal & Craig, 2001) usually rely on the work by Grandjean (1979) which outlines the main symptoms of mental fatigue including reduced alertness and impaired performance, implying that mental fatigue is defined and, consequently, measured indirectly, through the concepts of attention and performance. Empirical evidence from the rail context also supports this view: Smith and Smith stated that “Fatigue was no longer a significant predictor of performance when [other] variables were included” (2017, 257). However, attention is a critical safety capability directly impaired by fatigue (Williamson et al., 2011), and can be measured more objectively (see Towey et al., 2019). A challenge arising from this will be how to measure it *in real time*.

The second argument concerns the biophysiological underpinnings of fatigue and attention. Since fatigue is usually operationalized through concepts of attention or vigilance (i.e. sustained attention), research should focus on finding the physiological proxies of attention lapses, rather than fatigue. Attention is a cognitive concept with a clear biological substrate in the brain (for detailed overview of attentional networks see Peteresen & Posner, 2012). The closest proximity to measuring its activity is through EEG, which provides a more direct insight into it compared to other neuroimaging methods (Parasuraman et al, 2008). Since retina is considered to be ‘an approachable part of the brain’ (Dowling, 1987), we could expect eye-tracking measures to be the next best proxy of brain activity involved in attentional processes. The further away from the brain the signal recordings are, the weaker the link to attention and performance measures can be expected. This is well illustrated with the following empirical finding. An attempt was made to monitor attentional states with a wrist-wearable electrodermal potential (EDP), supposed to classify EEG states without the use of EEG cap (Mehler et al., 2018). Unlike HR data that was simultaneously measured, EDP-based attention metrics were not statistically significant in distinguishing between conditions varying in the level of attentional demand.

The third argument targets the core of the problem. As mentioned above, many studies interchangeably use the terms ‘fatigue’, ‘drowsiness’, and ‘sleepiness’. This not only brings conceptual confusion, but more importantly, addresses the issue when it is practically too late. If the real-time monitoring equipment provides feedback to the drivers only when they are drowsy or sleepy, the intervention can be ill-timed. It would be more useful if the monitoring system could recognize the onset of attention decline, which can be reversed.

Finally, even when drivers are perfectly rested and not experiencing any symptom of fatigue, they can still be distracted. Distractions could be internal, such as thoughts regarding personal-life matters, or external, such as highly salient stimuli in the surroundings; and can lead to performance decrements. A train driver’s job “requires *a* high level of concentration and alertness when it comes to signals, information, the tracks and his immediate environment” (Kecklund et al., 1999, 5). In addition to this, sustained attention was identified as a human factor that mostly contributed to all types of railway incidents, especially the inattentiveness to signals (Edkins & Pollock, 1997). Thus, a measure of driver attention could capture a wider range of instances when a driver’s performance is reduced, compared to a measure of fatigue.

5. General discussion and conclusion

This paper has critically examined the current state of the art in real-time fatigue assessment that can be applied to monitor train drivers’ fatigue. Several groups of methods for continuous and objective assessment of fatigue have been identified in the literature. Measures assessed as inapplicable to train driving were excluded. Further structured literature research was confined to three groups of physiological measures: EEG, eye-tracking, and heart-rate monitoring. Main variables that might be used to assess train driver fatigue are presented and analysed respectively. A summary table of key findings and issues in physiological assessment of fatigue across these measures and specific variables is provided.

Conclusions from this overview can be summarized as follows: based on the current state of the art available in literature, the single best fatigue measure could not be identified. Despite the constant advancements in driver-monitoring technology, there are no clear indications of solving the real-time fatigue detection issue that can be applied to train drivers. This seems to be more tied to the problem of fatigue conceptualization than to the reliability of available technology. Fatigue, or rather, mental fatigue, is often not clearly differentiated from sleepiness and sleep onset. In an attempt to measure fatigue, most of the reviewed studies rely on the concepts of attention or vigilance, and performance. Thus, this paper has proposed to redefine the problem of fatigue monitoring to that of attention monitoring. Four arguments were presented in detail to support this proposal, summarized as follows: 1) As opposed to fatigue, attention is a concept that can be measured objectively; 2) Attention has clear anatomical underpinnings in the brain, which can be monitored with EEG and eye-tracking; 3) Detecting and predicting attention lapses would bring a timelier solution compared to detecting fatigue (i.e. when it is already too late to intervene); and 4) Inattentiveness can occur even in the absence of fatigue, which covers more situations relevant to driving context.

More research endeavour is needed to determine the best way to measure and monitor attention levels of train drivers and explore if it is possible to do so unobtrusively and in real time. This paper can be useful for such future studies as it provides a critical overview of specific physiological variables that could be investigated in more depth in the context of monitoring train drivers' alertness and fatigue. It has also indicated several problems both in literature and practice that need to be addressed in future studies aiming to improve railway safety.

Acknowledgement

This work was conducted with the financial support of the Science Foundation Ireland Centre for Research Training in Digitally-Enhanced Reality (d-real) under Grant No. 18/CRT/6224.

References

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602-607.

Branton, P. (1979). Investigations into the skills of train-driving. *Ergonomics*, *22*, 155-164.

Bodala, I. P., Abbasi, N. I., Sun, Y., Bezerianos, A., Al-Nashash, H., & Thakor, N. V. (2017, July). Measuring vigilance decrement using computer vision assisted eye tracking in dynamic naturalistic environments. In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 2478-2481). IEEE.

Bodala, Indu P., et al. (2016). EEG and eye tracking demonstrate vigilance enhancement with challenge integration." Frontiers in human neuroscience 10 (2016): 273.

Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: an ERP study. Cognitive brain research, 25(1), 107-116.

Caffier, P. P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European journal of applied physiology*, *89*(3), 319-325.

Craig, A., Tran, Y., Wijesuriya, N., & Nguyen, H. (2012). Regional brain wave activity changes associated with fatigue. *Psychophysiology*, *49*(4), 574-582.

Crowley, K., & Balfe, N. (2018, July). Investigation of train driver physiological responses. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference 32* (pp. 1-9).

Di Stasi, L. L., Catena, A., Canas, J. J., Macknik, S. L., & Martinez-Conde, S. (2013). Saccadic velocity as an arousal index in naturalistic tasks. *Neuroscience & Biobehavioral Reviews*, *37*(5), 968-975.

Dowling, J. E. (1987). *The retina: an approachable part of the brain*. Harvard University Press.

Edkins, G. D., & Pollock, C. M. (1997). The influence of sustained attention on railway accidents. *Accident Analysis & Prevention*, *29*(4), 533-539.

ERA. 2020. *Report on Railway Safety and Interoperability in the EU*. [online] Luxembourg: Publications Office of the European Union, p.27.

Faber, L. G., Maurits, N. M., & Lorist, M. M. (2012). Mental fatigue affects visual selective attention. *PloS one*, *7*(10), e48073.

Fan, J., & Smith, A. P. (2020). Effects of occupational fatigue on cognitive performance of staff from a train operating company: a field study. *Frontiers in psychology*, *11*, 2366.

Filtness, A., and Naweed, A. (2017). Causes, consequences and countermeasures to driver fatigue in the rail industry: The train driver perspective. *Applied Ergonomics*, *60*, 12-21.

Fujiwara, K., Abe, E., Kamata, K., Nakayama, C., Suzuki, Y., Yamakawa, T., ... & Kadotani, H. (2018). Heart rate variability-based driver drowsiness detection and its validation with EEG. *IEEE Transactions on Biomedical Engineering*, *66*(6), 1769-1778.

Gramann, K., & Plank, M. (2019). The use of electroencephalography in neuroergonomics. In *Neuroergonomics* (pp. 11-15). Academic Press.

Grandjean, E. (1979). Fatigue in industry. *Occupational and Environmental Medicine*, *36*(3), 175-186.

Guarda, L., Astorga, N., Droguett, E., Moura, M., & Ramos, M. (2018). Drowsiness detection using electroencephalography signals: A deep learning based model. Proceedings of the Probabilistic Safety Assessment and Management PSAM, Los Angeles, CA, USA, 16.

Hamilton, W.I., and Clarke, T. (2005). Driver performance modelling and its practical application to railway safety. *Applied Ergonomics*, *36*, 661-670.

Hopstaken, J. F., van der Linden, D., Bakker, A. B., Kompier, M. A., & 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.

Horrey, W.J., Noy, Y.I, Folkard, S., Popkin, S.M., Howarth, H.D., and Courtney, T.K. (2011). Research needs and opportunities for reducing the adverse safety consequences of fatigue. *Accident Analysis and Prevention*, *43*, 591-594.

Hsieh, C. S., & Tai, C. C. (2013). An improved and portable eye-blink duration detection system to warn of driver fatigue. *Instrumentation Science & Technology*, *41*(5), 429-444.

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.

Jap, B. T., Lal, S., & Fischer, P. (2011). Comparing combinations of EEG activity in train drivers during monotonous driving. *Expert Systems with Applications*, *38*(1), 996-1003.

Jap, B. T., Lal, S., Fischer, P., & Bekiaris, E. (2007, August). Using spectral analysis to extract frequency components from electroencephalography: application for fatigue countermeasure in train drivers. In *The 2nd International Conference on Wireless Broadband and Ultra Wideband Communications (AusWireless 2007)* (pp. 13-13). IEEE.

Kecklund, G., Åkerstedt, T., Ingre, M., & Söderström, M. (1999). Train drivers' working conditions and their impact on safety, stress and sleepiness: a literature review, analyses of accidents and schedules. *National Institute for Psychosocial Factors and Health. Stress Research Report*, *288*.

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious?. *Perspectives on psychological science*, *7*(1), 18-27.

Lal, S. K., & Craig, A. (2002). Fatigue neurophysiology in professional and non-professional drivers. In *Proceedings of the Australasian road safety research, policing and education conference* (Vol. 6, No. 1, pp. 125-130). Monash University.

Lal, S. K., & Craig, A. (2001). A critical review of the psychophysiology of driver fatigue. *Biological psychology*, *55*(3), 173-194.

Lal, S. K., & Craig, A. (2000). Driver fatigue: Psychophysiological effects. In *International Conference on Fatigue and Transportation, 4th, 2000, Fremantle, Western Australia*.

LeBouef, T., Yaker, Z., & Whited, L. (2020). Physiology, Autonomic Nervous System. *StatPearls [Internet]*.

Lees, T., Chalmers, T., Burton, D., Zilberg, E., Penzel, T., Lal, S., & Lal, S. (2021). Electrophysiological Brain-Cardiac Coupling in Train Drivers during Monotonous Driving. *International Journal of Environmental Research and Public Health*, *18*(7), 3741.

Lees, T., Chalmers, T., Burton, D., Zilberg, E., Penzel, T., Lal, S., & Lal, S. (2018). Electroencephalography as a predictor of self-report fatigue/sleepiness during monotonous driving in train drivers. Physiological measurement, 39(10), 105012.

Liu, J., Zhang, C., & Zheng, C. (2010). EEG-based estimation of mental fatigue by using KPCA–HMM and complexity parameters. *Biomedical Signal Processing and Control*, *5*(2), 124-130.

Maffei, A., & Angrilli, A. (2018). Spontaneous eye blink rate: An index of dopaminergic component of sustained attention and fatigue. *International Journal of Psychophysiology*, *123*, 58-63.

Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., & Schwartz, P. J. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European heart journal*, *17*(3), 354-381.

Martins, R., & Carvalho, J. M. (2015). Eye blinking as an indicator of fatigue and mental load-a systematic review. *Occupational safety and hygiene III*, *10*.

McIntire, L. K., McKinley, R. A., Goodyear, C., & McIntire, J. P. (2014). Detection of vigilance performance using eye blinks. *Applied ergonomics*, *45*(2), 354-362.

McLeod, R.W., Walker, G.H., and Moray, N. (2005). Analysing and modelling train driver performance. *Applied Ergonomics, 36*, 671-680.

Mehler, B., Sawyer, B. D., McWilliams, T., & Reimer, B. (2018). An Applied Driving Evaluation of Electrodermal Potential as a Measurement of Attentional State. In *Neuroergonomics* (p. 321). Academic Press.

Morris, T. L., & Miller, J. C. (1996). Electrooculographic and performance indices of fatigue during simulated flight. *Biological psychology*, *42*(3), 343-360.

Naweed, A. (2014). Investigations into the skills of modern and traditional train driving. *Applied Ergonomics, 45*, 462-470.

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual review of neuroscience*, *35*, 73-89.

Peysakhovich, V., Dehais, F., & Causse, M. (2015). Pupil diameter as a measure of cognitive load during auditory-visual interference in a simple piloting task. *Procedia Manufacturing*, *3*, 5199-5205.

Sahayadhas, A., Sundaraj, K., & Murugappan, M. (2012). Detecting driver drowsiness based on sensors: a review. *Sensors*, *12*(12), 16937-16953.

Sangiorgio, V., Mangini, A. M., & Precchiazzi, I. (2020). A new index to evaluate the safety performance level of railway transportation systems. *Safety science*, *131*, 104921.

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in public health*, *5*, 258.

Shaffer, F., McCraty, R., & Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. *Frontiers in psychology*, *5*, 1040.

Shen, K. Q., Li, X. P., Ong, C. J., Shao, S. Y., & Wilder-Smith, E. P. (2008). EEG-based mental fatigue measurement using multi-class support vector machines with confidence estimate. *Clinical neurophysiology*, *119*(7), 1524-1533.

Shrestha, S., & Owens, J. W. (2008). Eye movement patterns on single and dual-column web pages. Usability News, 10(1), 0-6.

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(6), 679-692.

Smith, A. P., & Smith, H. N. (2017, June). Workload, fatigue and performance in the rail industry. In *International Symposium on Human Mental Workload: Models and Applications* (pp. 251-263). Springer, Cham.

Stern, J. A., Boyer, D., Schroeder, D., Touchstone, M., & Stoliarov, N. (1994). *Blinks, Saccades, and Fixation Pauses During Vigilance Task Performance: 1. Time on Task*. Washington University Seattle Dept Of Psychology.

Straeter, O. (2020, December). An Objective Measure for Detecting Workload and Errors in Monotone, Repetitive or Fatigue-Causing Environments Using Pupil Variation. In *International Symposium on Human Mental Workload: Models and Applications* (pp. 3-12). Springer, Cham.

Towey, G.E., Fabio, R.A., & Caprì, T. (2019). Measurement of attention. In Caprì, T., Fabio, R.A., Towey, G.E., & Antonietti, A. (Eds). Psychology Research Progress. New York: Nova Science Publisher. ISBN: 978-1-53615-441-2

Trutschel, U., Sirois, B., Sommer, D., Golz, M., & Edwards, D. (2011). PERCLOS: An alertness measure of the past.

Wang, H., Zhou, L. B., & Ying, Y. (2010, June). A novel approach for real time eye state detection in fatigue awareness system. In *2010 IEEE Conference on Robotics, Automation and Mechatronics* (pp. 528-532). IEEE.

Williamson, A., Lombardi, D.A., Folkard, S., Strutts, J., Courtney, T.K., and Connor, J. (2011). The link between fatigue and safety. *Accident Analysis and Prevention, 43*, 498-515.

Young, M.S. and Steel, T. (2016). Rail worker fatigue: Identification, management and countermeasures. *Journal of Rail and Rapid Transit, 231*, 1098-1106.

Yuzhong, L., Zhiqiang, L., Zhixin, Y., Hualiang, L., & Yali, S. (2020, October). Research Progress of Work Fatigue Detection Technology. In *2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT* (pp. 558-560). IEEE.