Countering Train Driver Fatigue in Automatic Train Operation

Niels BRANDENBURGER, Martin WITTKOWSKI and Anja NAUMANN

German Aerospace Centre, Germany

Abstract.

Extensive levels of train driver fatigue are associated with impaired performance. Increasing automation is expected to result in more driver fatigue. We assessed physiological and subjective fatigue in 26 train drivers in manual and automated driving mode in a two hour simulator study. Results show a significant progression of fatigue over time across measures, but do not reveal automation effects on fatigue. The results, possible conclusions, and a driver-centered approach to Automatic Train Operation (ATO) are discussed.

Keywords. Automatic train operation; Train driver; Fatigue; EEG

1. Introduction

Fatigue has major implications for transportation system safety and is believed to present a major hazard in the transportation industry. Therefore, investigating the psychophysiological links to fatigue could enhance the understanding and management of fatigue in the transportation industry (Lal & Graig, 2002). The term fatigue is widely used to indicate the influence of long working periods, reduced rest, and being unable to sustain a certain level of task performance (Dinges, 1995; see Lal & Graig, 2002). Fatigue generally impairs human efficiency when individuals continue working after they have come aware of their fatigue state (Lal & Graig, 2002). Focusing on the performance aspect of fatigue, driver fatigue has been defined as a state of reduced mental alertness that impairs performance during a range of cognitive and psychomotor tasks including driving (see Lal & Graig, 2002). In the railway domain, increasing numbers of technical systems supporting the train driver through automatic train protection (ATP) and ultimately automatic train operation (ATO) raise questions about established effects on train driver fatigue. Comprehensive research has been conducted on fatigue and the corresponding psychophysiological links across transportation domains. Aggregated results concerning the effects of fatigue on car driving reveal slower reaction times, deterioration of steering performance and headways, dangerous individual compensatory strategies and loss of awareness (DaCoTA, 2012). Factors fostering fatigue include the length of a journey, monotonous driving situations, time of the day as well as irregular driving schedules (Lal & Graig, 2002). A study with truck drivers reported cortical deactivation and increased sleepiness at the end of an all-night driving shift (Kecklund & Åkerstedt, 1993). Likewise, studies with nonprofessional drivers demonstrated cortical deactivation in response to continuous and monotonous driving (Brookhuis & de Waard, 1993). Apart from external factors, the driving task itself may lead to decreased physiological arousal, slowed sensorimotor...
functions, and impaired information processing; effectively diminishing a driver’s ability to respond appropriately to regular and emergency situations (Mascord & Heath, 1992; Lal & Graig, 2002).

Measures of fatigue stem from subjective, objective performance-based and physiological categories, because of the difficulty to grasp the multifaceted concept of fatigue with a single measure. This emphasizes the need to tap into at least two categories of measures. Focusing on the physiological category, promising results linked to fatigue stem from Electroencephalography (EEG) (DaCoTA, 2012). EEG seems to be one of the most predictive and reliable techniques for detecting changes in alertness and fatigue related vigilance decrements (see Lal & Graig, 2002; Jap, Fischer & Bekiaris, 2007; Borghini et al., 2012). Lal and Graig (2002) showed fatigue to be associated with significant increases in slow delta and theta band brain activity and decreases in fast brain activity (beta and gamma band); a finding replicated in several studies (e.g. Ridwan et al., 2009; Borghini et al., 2012). Heart rate is another physiological indicator linked to fatigue. A decrease of the heart rate (HR) and an increase of the heart rate variability (HRV) can be interpreted as an indicator for fatigue (Wierwille & Muto, 1981; Hefner et al., 2009). As Lal & Graig (2002) report in their car driving simulator study mentioned earlier, heart rate was significantly lower after the driving task. Within the subjective category of tools to measure fatigue, the Karolinska Sleepiness Scale (KSS) has been reported to be sensitive to the subjective aspect of fatigue also described as sleepiness (DaCoTA, 2012).

1.1 Train driver fatigue

Train driver fatigue has long been recognized as a serious issue in railway operations (Grant, 1971). Jap, Fischer and Bekiaris (2007) state that “most train accidents can be attributed to fatigue”. In train driving, fatigue occurs rather seldom as a consequence of information overload, but is rather caused by underload (Grant, 1971). The negative effect of train driver underload on fatigue occurs separately from other fatigue inducing working conditions as for example shiftwork (e.g. Leutner & Debus, 1995; Härmä et al., 2002; DaCoTA, 2012), working hours (Buck & Lamonde, 1993) or monotony (Dunn & Williamson, 2011; Dorrian et al., 2006; Stein & Naumann, 2016). Fatigue has a negative effect on performance (Dunn & Williamson, 2011). Dorrian et al. (2007) reported especially errors of omission and seizing manual interaction with the train to be associated with fatigued train drivers. Furthermore the train divers’ ability to react to unpredictable events deteriorates with increasing fatigue (Schmitz, 2013). Substantial assessment of train driver fatigue through physiological indicators is scarce, although some efforts have been made. Torsvall and Åkerstedt (1987) recorded the EEG of train drivers and found activity in the alpha, delta and theta band to be sensitive to fatigue. In a similar experimental setup Ridwan et al. (2009) reported activity in the lower frequency delta and theta bands to rise with increasing fatigue, replicating earlier results by Lal and Craig (2002). In line with this notion Jap, Fischer and Bekiaris (2007) reported decreasing high frequency activity in the beta band to go along with fatigue. Building upon these insights into power spectral changes in the EEG, Michel et al. (1992) started to localize changing frequency bands within the brain of fatigued train drivers, giving rise to rather current approaches by Jap, Fischer and Lal (2011) to develop online algorithms for fatigue detection. Apart from EEG, fatigue indicators based on heart rate have received attention in the railway domain (e.g. Grant, 1971). In a recent study, Stein and Naumann (2016) conducted a monotonous train simulator study showing that...
increased fatigue was associated with a decrease in train driver performance, a decrease of the heart rate (HR) and an increase of the heart rate variability (HRV) after the monotonous ride.

1.2 Effects of automation on train driver fatigue

In the domain of train driving, a shift from physical work to mental information processing and continuous monitoring has taken place over the last decades (Naweed, 2013; Brandenburger et al., 2016). Rising levels of automation characterize modern train protection systems as for example the European Rail Traffic Management System (ERTMS; European Railway Agency, 2007). ERTMS includes advisory systems for speed and optimal traction deployment. Automatic train operation drastically changes the train driver’s tasks and required skills (Brandenburger et al., 2016). These changes in the task environment towards higher levels of automation might increase the train driver’s vulnerability to fatigue (Brandenburger et al., 2016). Human inefficiency in monitoring in general and specifically monitoring automated systems for infrequent failures has long been established in the literature (e.g. Molloy & Parasuraman, 1996; Onnasch et al., 2014). Molloy and Parasuraman (1996) reported decreased human monitoring performance when a system was automated in contrast to manual control. A meta-analysis provided by Onnasch et al. (2014) concluded that high degrees of automation had a negative impact on human performance in response to system failure as well as negative consequences for situation awareness. Spring et al. (2008) studied train driving in four levels of automation and replicated the conclusion of Onnasch et al. (2014) in the railway domain. Their study reveals that high levels of automation are associated with the lowest performance in detection of a rail signal failure in a train driving simulator. Those railway specific results were however based on a student population and did not include specific measures for fatigue. Therefore in this study we aimed to investigate two consecutive research questions: a) whether driving manually under ERTMS in-cabin signaling significantly increases fatigue within a relatively short two hour drive and b) whether continuous supervision of automatic train operation (ATO over ETCS) results in more fatigued train drivers in the same period of time. Therefore the following hypotheses concerning the effect of time (H1) and Automation (H2) will be tested on the basis of Ridwan et al. (2009), Lal and Craig (2002), Jap Fischer and Bekiaris (2007), Naumann and Stein (2016) and Brandenburger et al. (2016):

Representing the effect of time on fatigue within subjects, we expect

- \( H1.1: \) the power in the delta band to increase over time
- \( H1.2: \) the power in the theta band to increase over time
- \( H1.3: \) the power in the beta band to decrease over time
- \( H1.4: \) the heart rate to decrease over time
- \( H1.5: \) the heart rate variability to increase over time
- \( H1.6: \) the KSS score to increase over time.

Representing the effect of automation on fatigue between subjects, we expect, compared to the manual condition,

- \( H2.1: \) the power in the delta band to be higher in the automatic speed control condition
- \( H2.2: \) the power in the theta band to be higher in the automatic speed control condition
H2.3: the power in the beta band to be lower in the automatic speed control condition
H2.4: the heart rate to be lower in the automatic speed control condition
H2.5: the heart rate variability to be higher in the automatic speed control condition
H2.6: the KSS score to be higher in the automatic speed control condition.

2. Methods

2.1 Participants

The sample for this study consisted of 26 male active train drivers aged 21 to 56 (M = 36.53, SD = 10.92) from Germany. The professional background was in freight transport (four participants) and in passenger transport (22 participants). Occupational experience ranged from 1 to 37 years of experience (M = 14.07, SD = 10.85). None of the participants had prior experience with any form of automatic train operation. Participants reported to be free of chronic cardiac and neurophysiological conditions and absent from regular intake of medicaments. All participants were reimbursed with 30 € for their participation.

2.2 Experimental Setup and Measures

A between-subject design incorporating one independent variable “speed control” with two levels: "manual speed control" and “automatic speed control” was chosen to investigate the above mentioned research questions in the high fidelity railway simulator “RailSET” (Brandenburger, Stamer & Naumann, 2017) in Brunswick, Germany. The simulator features ERTMS baseline two functionality and an automatic speed control system that can be activated and deactivated by the driver in real-time. To enable the analysis of the dependent measures over time, the total driving time of approx. two hours was split into three separate blocks of approx. 40 minutes each by ordering the participants to stop shortly at a station. The route was characterized by a monotonous high speed track mainly secluded by sound-proofing walls. ETCS Level 2 in-cabin signaling without line sight signaling was chosen as a train protection system. Automatic speed control was implemented to look like “ATO over ETCS” (European Railway Agency, 2016). Hence, the simulation environment was set up to resemble future high speed passenger transport. The following measures were used to assess the dependent variables: Physiological measures: EEG and ECG were recorded by a Poly Data Recorder (PD3), which was developed by DLR e.V. for ambulatory measurements. The EEG was derived from C3/A2 and C4/A1 according to the international 10-20 system. Automatic speed control was implemented to look like “ATO over ETCS” (European Railway Agency, 2016). Hence, the simulation environment was set up to resemble future high speed passenger transport. The following measures were used to assess the dependent variables: Physiological measures: EEG and ECG were recorded by a Poly Data Recorder (PD3), which was developed by DLR e.V. for ambulatory measurements. The EEG was derived from C3/A2 and C4/A1 according to the international 10-20 system. The electrocardiographic lead (ECG) in is a variant of Einthoven’s lead, where a placement diagonally to the chest is recommended. EEG and EOG were sampled with 256 Hz, the ECG with 1024 Hz. In the selection of electrode positions, attention was paid so as not to interfere with the driving tasks. The signals were filtered online by hardware to minimize artifacts (0.1 Hz to 80 Hz included; 50-Hz notch filter). After post-hoc artifact correction, the EEG signals were evaluated according to Rechtschaffen & Kales (1968). Artifacts usually arising from body movements, blinking and other muscle activity were identified through a two-step process, i.e. 1) an automated algorithm, and 2) visual inspection by an expert scorer. The EEG signals were analyzed by means of a fast Fourier transform (FFT) routine at a window length of 2 seconds, which corresponds to a frequency resolution of 0.5 Hz. EEG power densities within the following frequency bands were included in the analysis of fatigue: 0.5 Hz – 3.5 Hz (Delta), 4 Hz – 7.5 Hz (Theta), 8 Hz – 12.5 Hz
(Alpha), 13 Hz - 18 Hz (Beta), 19 Hz – 30 Hz (Gamma). Power densities were averaged separately within each of the three blocks. The inter-beat-interval and beat-to-beat heart rate was automatically extracted from the ECG signal and used for heart-rate variability (HRV) analysis. **Subjective measures:** Subjective fatigue was assessed using a paper-based version of the Karolinska Sleepiness Scale (KSS; e.g. Kaida et al., 2006). Participants indicated their subjective level of fatigue on a 9-point Likert scale ranging from “extremely alert” to “very sleepy, great effort to keep awake, fighting sleep” multiple times over the course of the study. Additionally, the Epworth Sleepiness Scale (ESS; Murray, 1993) was administered prior to the simulator study to assess a baseline of subjective fatigue. Participants indicate the chance of dozing in eight imaginary situations on a 4-point Likert scale ranging from “would NEVER doze” to “High chance of dozing”. Further materials included demographic questionnaire, informed consent, handout about ETCS and debriefing form. The ETCS handout differed for manual and automatic speed control conditions.

2.3 Procedure

Upon arrival the participants filled in the demographic questionnaire, the informed consent and the ESS. Afterwards, they read the ETCS handout and the physiological equipment was set up and calibrated. Next, they filled in the KSS and started driving for three consecutive blocks of approx. 40 minutes each. Between blocks the participants shortly stopped at a station and filled in the KSS. After finishing the route, participants filled in the KSS for the last time, the equipment was detached and participants were debriefed.

3. Results

Prior to analyses related to the hypotheses, random confounding effects of caffeine intake and daytime between the speed control groups were investigated in all dependent variables and found not to be present in the dataset. The baseline fatigue assessed through the ESS revealed no confounding differences between the two groups of speed control. Additionally, all results that are based on repeated measures ANOVA models were corrected according to the Greenhouse-Geisser correction to fulfill the assumption of sphericity. Furthermore, no interactions between block and speed control were present in any of the models we report.

**Table 1. Summary of effects for Block and Speed control on the dependent variables. Levels of significance are as follows: * = <.1; ** = <.05; *** = <.01.**

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>Block Effects</th>
<th>Speed control Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>p</td>
</tr>
<tr>
<td>Delta power</td>
<td>10.52</td>
<td>.001***</td>
</tr>
<tr>
<td>Theta power</td>
<td>5.65</td>
<td>.007***</td>
</tr>
<tr>
<td>Beta power</td>
<td>4.34</td>
<td>.019**</td>
</tr>
<tr>
<td>Heart rate</td>
<td>23.25</td>
<td>.000***</td>
</tr>
<tr>
<td>Heart rate variability</td>
<td>20.97</td>
<td>.000***</td>
</tr>
<tr>
<td>KSS</td>
<td>15.79</td>
<td>.000***</td>
</tr>
</tbody>
</table>
For each dependent variable, a repeated measures ANOVA model was set up and the effects of the within-subject variable block, representing the time-on-task effect, and the between-subject variable speed control were tested for significance at an alpha level of .05.

Table 1 summarizes the effects in the gathered data of the 26 participants. Concerning the effect of the driving block, the results show a clear effect of time across all dependent variables. Delta and Theta power significantly increased over the three blocks, confirming hypotheses H1.1 and H1.2 (Figure 1). Conversely, power in the beta band significantly dropped over the three blocks, which is in line with H1.3.

Figure 2. Changes in power density over the three consecutive driving blocks for all frequency bands. 0.5 Hz – 3.5 Hz (Delta), 4 Hz – 7.5 Hz (Theta), 8 Hz – 12.5 Hz (Alpha), 13 Hz – 18 Hz (Beta), 19 Hz – 30 Hz (Gamma).

Also the heart rate dropped over time and the heart rate variability increased, both at a probability well below the alpha level and therefore confirming H1.4 and H1.5. The KSS scores representing subjectively perceived fatigue also increased significantly over time, confirming H1.6, although there is an interesting pattern in the fourth and last measurement (Figure 2), revealing a slight drop of the KSS scores. Between-subject effects of speed control are less prominent in the data. None of the models for all dependent variables showed a significant effect of speed control. Therefore H2.1 - H2.6 need to be dismissed based on the results of this study, although there is a trend in theta power band...
Participants in the automatic speed control groups tended to have higher power in the theta power band than participants in the manual speed control group. In contrast to H2.3 and H2.4, average beta power and the average heart rate were higher in the automatic speed control group.

Figure 2. Summary of Karolinska Sleepiness Scale scores before (t1), during (t2 and t3) and after (t4) the study for all participants.

4. Discussion and Conclusion

The goal of this study was to investigate whether a) driving manually under ERTMS in-cabin signaling significantly increases fatigue within a relatively short two hour drive and b) whether continuous supervision of automatic train operation results in more fatigued train drivers in the same period of time. Concerning the first research question, the study provides clear evidence that driving under ERTMS in-cabin signaling for approx. two hours exerts a detrimental effect on alertness, resulting in subjectively and physiologically fatigue of train drivers. All six dependent variables show highly significant increases in fatigue levels in all participants after two hours. This is in line with results from Naumann, Wörle & Dietsch (2016), who reported more attentional resources to be directed onto the displays within the cabin as levels of automation rise. Under ERTMS and more specifically with ETCS Level 2 without line sight signaling, the in-cabin displays are the main sources of information. Drivers are required to constantly monitor them. This continuous monitoring creates an underloading task environment (Grant, 1971) and the current results show that this fosters fatigue within a relatively short amount of time; a position forwarded by Brandenburger et al. (2016) on the basis of earlier research by Molloy and Parasuraman (1996) as well as Onnasch et al. (2014). The results of the present study replicate findings on cortical deactivation of high frequency activity in response to
continuous monitoring reported in the context of car drivers (Brookhuis & De Waard, 1993) in a sample of train drivers. This raises questions about the effects of driving under ETCS Level 2 without line sight signaling on the alertness and ultimately the performance of the train driver during prolonged working hours. The effects of increasing automation represented by ATO over ETCS on the subjective and physiological fatigue levels are more complex to interpret. Mainly, we did not find significant differences between manual and automatic speed control groups in the current study. Yet interestingly, we found results in beta power and heart rate measures that seem to be opposite to H2.3 and H2.4. Train drivers driving with automatic speed control tended to have more high frequent brain activity and a higher heart rate than manually driving participants. As these participants did not have any prior experience with ATO over ETCS and potential confounding variable have been discarded, we assume an unintended novelty effect within this group. It can be assumed that being unfamiliar with ATO and rather suspiciously monitoring the automatic system without any reference on reliability caused the participants to stay more alert, resulting in a higher heart rate and more high frequency brain activity throughout the study. Therefore, the process of familiarization of the drivers with ATO systems needs to be attended very closely to avoid the pendulum to swing from distrust to over trust and phenomena such as complacency (Singh, Molloy & Parasuraman, 1993; Parasuraman & Manzey, 2010). However, based on the negative impact of driving under ERTMS on fatigue, we expect the consequences for fatigue to increase in real railway operations with shifts of seven and more hours in unsupervised working conditions. Therefore, future research should focus on skipping the automation level where continuous monitoring is the central human task in favor of higher grades of automation, where only situation specific supervision of automatic trains is required (Brandenburger et al., 2017). Designing a task environment placed, for example, in a control center that enables the train driver to supervise and control different vehicles on the track on demand, may unleash the full potential of ATO on the main line in terms of efficiency as well as provide the future train driver with a less monotonous task environment that supports optimal human.

References


DaCoTA. (2012). Fatigue: Deliverable 4.8h of the EC FP7 Project DaCoTA. Retrieved from www.dacota-project.eu


